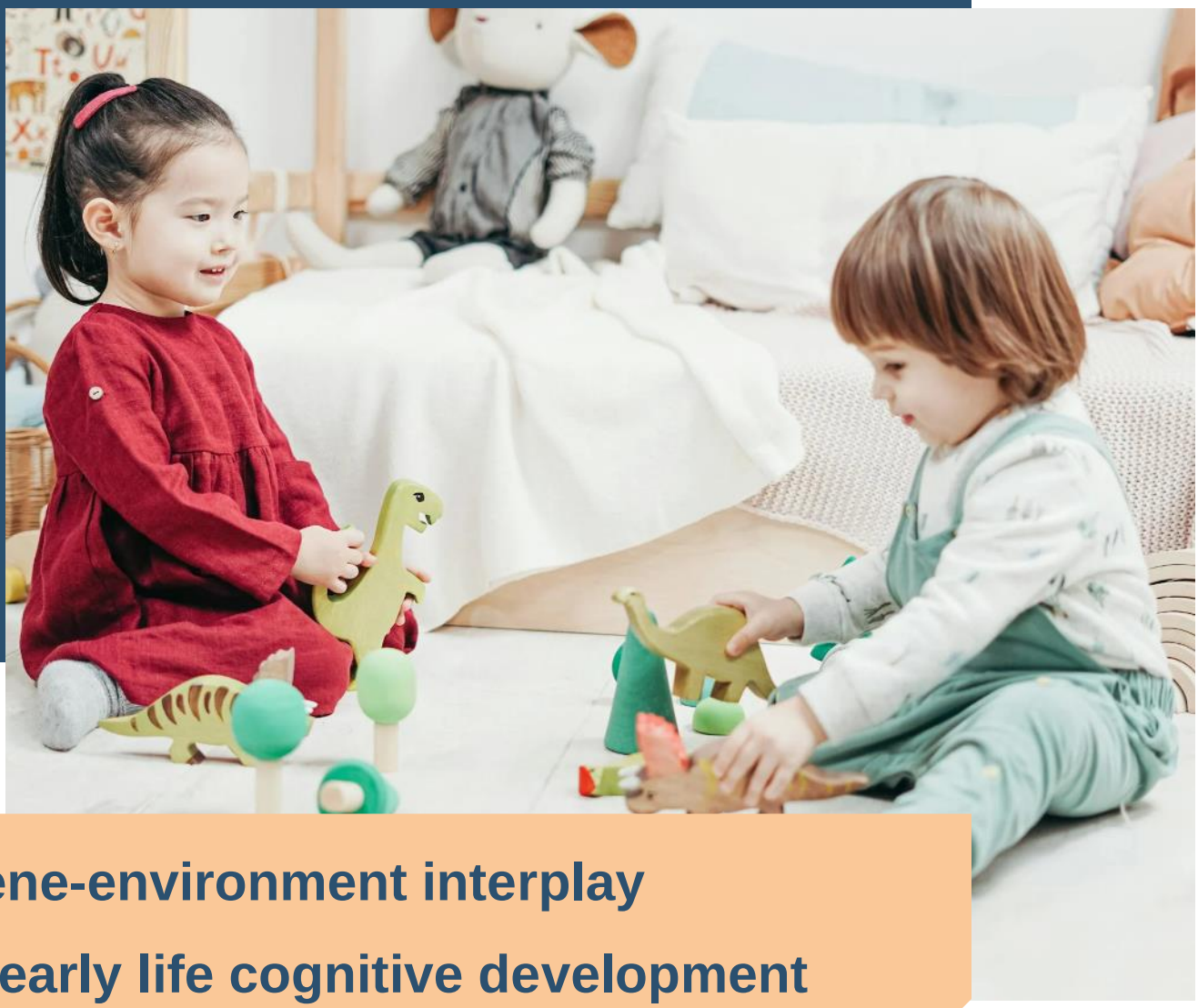


# Gene-environment interplay in early life cognitive development

Final report

Alexandra Starr, Florence Oxley, & Sophie von Stumm

Hungry Mind Lab, Department of Education, University of York, United Kingdom

March 2024

## Acknowledgements

We are grateful for the ongoing contribution of the participants in the Twins Early Development
Study (TEDS) and their families. TEDS is supported by the UK Medical Research Council (MR/V012878/1 and previously MR/M021475/1), with additional support from the US National Institutes of Health (AG046938). We thank our Advisory Board members for their invaluable advice and comments throughout this project. This project was supported by the Nuffield Foundation award (EDO/44110) to SvS on “The gene-environment interplay in early life cognitive development”.

This project has been funded by the Nuffield Foundation.
The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org


Citation:

Starr, A., Oxley, F., & von Stumm, S. (2024). Gene-environment interplay in early life cognitive development (Final report). Nuffield Foundation.

# Table of Contents

## Executive Summary

### *Introduction*

Cognitive ability – our capacity for reasoning, thinking, and learning – predicts a wide range of important life outcomes, including access to education, work, and civic life. Likewise, social-emotional development – the ability to express, understand, and regulate our own and others' emotions, and to form and maintain relationships with others – predicts our life outcomes. Children begin to show differences in cognitive ability and social-emotional behaviour during the early years of life and these differences magnify across the lifespan. Children's differences in cognitive and social-emotional development result from the interplay between their inherited genetic differences and their environmental experiences. The current project set out to explore one specific aspect of this interplay: We sought to identify which environmental experiences interacted with children's genetic propensities in the development of their early life cognitive abilities and social-emotional traits. The DNA revolution made it possible to identify some of the specific DNA differences, passed on from parents to children, that are linked to different traits. These DNA differences can be aggregated to create so-called polygenic scores – values that indicate how likely a person is to, for example, be emotionally stable or do well at school. This project constitutes the first systematic study of gene-environment interactions in early life that used polygenic scores for cognitive and social-emotional development and tested their interactions with a broad range of environmental factors.

### *Background and Aims*

Children's differences in early life cognitive abilities and social-emotional traits are driven by a complex interplay between biological and experience-dependent processes of development. According to behavioural genetics – the study of how our genetic differences influence how we think and behave – the gene-environment interplay is borne out through two key processes: *gene-environment correlation* and *gene-environment interaction*.

*Gene-environment correlation* (rGE) describes the way in which inherited genetic differences are associated or correlated with a person's environmental experiences and vice versa. Genes and environments are systematically linked, rather than co-occurring at random. There are three different types of rGE. *Passive rGE* occurs when a genetic propensity for a given trait, like reading aptitude and enjoyment, is passed from parent to child both genetically, via DNA, and

environmentally (e.g., by parents keeping books in the family home, which children may read). Continuing with this example, children who enjoy reading may evoke more or stronger encouragement to read from adults around them. This is termed *evocative rGE*. Finally, *active rGE* can occur when children who enjoy reading create their own opportunities to read more by visiting libraries or sourcing books themselves.

*Gene-environment interaction* (GxE) explains why individuals come to develop different traits through the mutual influence between genetic and environmental factors. For example, two siblings brought up in the same family home may show different traits because of underlying differences in their genetic makeup. Conversely, two individuals who share a common genetic propensity to a given trait or traits may show different outcomes if they have different environmental experiences. GxE imply that genetic propensities for a trait may be more or less predictive of developing the trait phenotypically depending on the environment. There are three principal theories on GxE. Continuing, again, with the examples of reading aptitude and enjoyment, the *diathesis-stress model* suggests that children who are less strongly genetically predisposed towards these traits may be at a greater disadvantage than children who are more strongly genetically predisposed towards these traits if exposed to literacy deprived environments, for example when they receive little encouragement to read. The *bio-ecological model* suggests that a high genetic propensity for reading aptitude and enjoyment may be phenotypically expressed in stimulating reading environments, but may be suppressed if a child is not given adequate exposure to stimulating reading environments. The *model of differential susceptibility* suggests that children with a lower genetic propensity for reading aptitude and enjoyment may benefit more than children with higher propensities to these traits from stimulating reading environment and, at the same time, may be more disadvantaged if they are not given stimulating reading environments. In summary, rGE describes how genetic tendencies and environmental experiences go hand-in-hand, whereas GxE describes the continuous ‘conversation’ that takes place between them throughout development. Both rGE and GxE offer helpful insights into how and why individual differences in complex psychological traits emerge.

Yet to date, no robust evidence of specific gene-environment interactions in the prediction of early life cognitive and social-emotional development has been found, although some studies reported GxE effects in the prediction of outcomes when children were older (e.g., Cheesman et al., 2022). To support children’s development effectively from as early as possible, building an empirical evidence-base of GxE in cognitive and social-emotional development is key.

Understanding which environmental factors and genetic propensities exert most influence on the expression of cognitive and social-emotional traits would allow researchers and practitioners to design and implement effective interventions that stimulate the development of these traits. The studies we conducted during this research project investigated relationships between polygenic scores for years spent in education, childhood ADHD, and neuroticism, and 39 types of environmental experience, which we divided into four categories. These categories were homelife and learning environment (e.g., sleep schedule, siblings, learning resources and toys), neighbourhood (housing and facilities, education, employment), childhood adversity (life events, parent and child health risks), and exposure to pollution (harmful chemical exposure).

### *A cautionary note on genetic propensities*

Genetically-informed research on psychological or behavioural traits is often met with scepticism and even hostility, and findings relating to heritability are frequently misinterpreted. A common misconception is that genetic influences mean that people's differences are 'fixed' or 'destined' – an idea also known as biological or genetic determinism. In reality, genetic influences neither determine destiny nor do they imply immutability. Rather, genetic and environmental factors both influence our thinking and our behaviour, each to a sometimes greater, sometimes lesser extent. If environments vary greatly in a population, for example when the opportunity to go to school is only available to some but not others, genetic influences will only explain a small proportion of children's differences in educational outcomes, like years spent in school or school grades. In this case, the effects of having access to a school versus having little or no access to school would have the much stronger impact on children's educational success. However, if environmental conditions are more equal for people, then genetics will account for a greater proportion of the differences seen between them. For example if schooling is available to, or even compulsory to a certain age for all children in a given population (i.e., when there is less difference between people's environmental experiences), then more of those children's differences in educational success may be caused by differences in their individual genetic propensities for learning.

DNA-based predictions for phenotypic traits like cognitive and social-emotional development reflect people's differences in their genetic propensities to think, feel, and behave in certain ways. For example, children with lower genetic propensities for cognitive ability might on average find concentrating on homework from school more difficult, might be less engaged with

lessons, and might forgo study sessions in favour of other activities more often than children with higher genetic propensities for cognitive ability. It may therefore take more effort, time, and educational support for children with lower genetic propensities for cognitive ability to achieve the same educational success as others with higher propensities, but they could nonetheless perform just as well or even better. At the other end of the spectrum, children with higher genetic propensities for cognitive ability may show greater interest and aptitude in studying at school and home, but this tendency may be over-ridden if a child suffers severe adverse experiences in school or their home learning environment. In other words, children's genetic propensities for cognitive ability and social-emotional traits describe tendencies or preferences but are not solely responsible for determining their developmental trajectories.

### *Methods*

We capitalised on data from identical and fraternal twins participating in the Twins Early Development Study (TEDS), a genetically informative, representative cohort study in the UK that comprises data from initially over 15,000 families with twins born between 1994 and 1996 in England and Wales. Over the course of this longitudinal study, twins' verbal and nonverbal cognitive development was assessed at the ages 2, 3, 4 years via standardised testing and extensive parent and self-reports. Children's social-emotional behaviour was assessed at ages 7, 9, and 12 using information from teacher reports about the extent to which children's showed problems with peer relationships, conduct, inattention/hyperactivity, and emotional regulation. Additionally, information on a wide range of environmental factors were assessed (e.g., parental behaviour, the family home environment, early life experiences, family adversity, neighbourhood characteristics and measures on environmental pollution).

Children's genetic propensities for cognitive and social-emotional outcomes were operationalised as polygenic scores that aggregated DNA variants associated with cognitive and social-emotional traits. To predict cognitive outcomes, we used polygenic scores for years spent in full-time education (Lee et al., 2018). Although these polygenic scores were aimed at capturing adults' educational attainment, they are also powerful predictors of individual differences in learning-related traits, such as school performance and intelligence (e.g., Allegrini et al., 2019). For prediction of social-emotional outcomes, we used polygenic scores for ADHD (Demontis et al., 2019) and neuroticism (Luciano et al., 2018), since these explain substantial variance in children's social-emotional behaviour (Gidziela et al., 2021).

To investigate effects of gene-environment interactions in early life development, we applied regression-type analyses to predict development from environmental and genetic measures as well as their interaction effects. We included environmental factors in our models in the form of single measures as well as composite scores. This allowed us to test how strongly children's environmental experiences predicted their cognitive ability and social-emotional traits. We differentiated different levels of measurement to model the environment, including individual environmental experiences (e.g., having books in the family home vs not) and sets of related environmental experiences (e.g., ratings of the home learning environment). As an alternative methodological approach, we investigated the underlying factors that drive associations between child development and environmental factors using a cross-lagged twin model. For the purpose of developing recommendations around statistical power in future GxE research, we conducted simulation studies to estimate the sample sizes that are required to robustly identify interaction effects.

### *Findings*

The findings from our studies on GxE in children's early life development conducted over the course of this project did not reveal robust GxE effects on children's cognitive ability or their social-emotional development. The broad range of environmental measures included in our analyses predicted up to 20% of the variance in children's developmental outcomes. By comparison, children's genetic propensities operationalised as polygenic scores explained independently only up to 2% of the variance in their cognitive and social-emotional developmental outcomes. These findings suggest that, in early life, environmental factors are stronger predictors of developmental differences than DNA-based predictors. We hasten to add that these differences in prediction strengths do not allow inferring conclusions about the importance of environmental factors relative to that of genetic influences on children's early life development. Rather they reflect the amount of variance in children's cognitive and social-emotional traits that can be attributed to genetic and environmental factors using currently available methods. The environmental and genetic prediction estimates may change in the future as we learn more about how our individual differences unfold over time.

We found direct effects of environments and polygenic scores, but no significant interaction effects emerged that explained significantly more variance in children's development. However, we do not interpret our findings to imply that there are no GxE in children's early life

development. Instead, we suspect that our results are false negatives, which may have emerged for three main reasons. First, our studies might not have had sufficient statistical power to robustly identify the potentially miniscule and complex GxE effects on children's cognitive and social-emotional development. To test this hypothesis, we conducted novel simulation studies on statistical power and found that sample sizes of over N = 70,000 would be needed to detect interactions of realistic effect sizes. For this project data were available from only about 7,000 children. Second, the measures we used to operationalise environments and genetic propensities may have been unsuitable for identifying effects at the level of observation at which GxE occur. Since both our genetic and environmental measures are composites comprising information at a comparatively coarse level, we cannot rule out that GxE effects may exist at a more fine-grained level, that is, at the level of single genetic variants and narrow environmental factors. Lastly, we investigated alternative modelling approaches to the gene-environment interplay to test if children's cognitive development 'caused' the cognitive stimulation they received in early life for genetic reasons, or if their cognitive development was 'caused' by their environmental experiences in the family home. We found bidirectional effects, due to common genetic and shared environmental factors which suggested that children are not passive recipients of the environments that they are exposed to. Instead, they actively select, shape, and create the environmental experiences that match their genetic propensities.

## Overview of peer-reviewed scientific publications

Oginni, O., Starr, A., & von Stumm, S. (2024). Do children cause the cognitive stimulation they receive? Modelling the direction of causality. *PsyArXiv*. https://doi.org/10.31234/osf.io/pqf78

Oxley, F. A. R., Wilding, K., & von Stumm, S. (in progress). DNA & IQ: Big deal or much ado about nothing? Preprint: https://osf.io/63zmr/

Plomin, R., Gidziela, A., Malanchini, M., & von Stumm, S. (2022). Gene–environment interaction using polygenic scores: Do polygenic scores for psychopathology moderate predictions from environmental risk to behavior problems? *Development and Psychopathology*, 1-11. https://doi.org/10.1017/S0954579422000931

Plomin, R., & von Stumm, S. (2022). Polygenic scores: prediction versus explanation. *Molecular Psychiatry, 27,* 49–52, https://doi.org/10.1038/s41380-021-01348-y.

von Stumm, S., & Plomin, R. (2021). Using DNA to predict intelligence. *Intelligence*, 86, 101530. https://doi.org/10.1016/j.intell.2021.101530

(Awarded the Excellence in Research Award, Mensa Education & Research Foundation)

von Stumm, S. (2022). Early childhood inequalities: The rocky path from observation to action. *IFS Deaton Review of Inequalities*. https://ifs.org.uk/inequality/early-childhoodinequalities-the-rocky-path-from-observation-to-action

von Stumm, S., Kandaswamy, R., & Maxwell, J. (2023). Gene-environment interplay in early life cognitive development. Intelligence, 98, 101748. https://doi.org/10.1016/j.intell.2023.101748

von Stumm, S. & Nancarrow, A. (2023). New methods, persistent issues, and one solution: Gene-environment interaction studies of childhood cognitive development. *PsyArXiv*. https://osf.io/preprints/socarxiv/y28h6

Wilding, K., Wright, M., & von Stumm, S. (2023). Using DNA to predict Educational Success: A Meta-Analytic review. *PsyArXiv*. https://doi.org/10.31234/osf.io/f7vt9

## Glossary of terms common in behavioural genetic research

(Adapted from Wilding et al., 2023)

| Term | Definition |
|---|---|
| Causality and prediction | *Causality* describes a relationship between two events, whereby one event causes the other as a direct result e.g., when clouds become saturated, this causes it to rain.<br>By contrast, *prediction* describes a relationship whereby two or more events are linked in such a way that if one event takes places, the other event is also very likely to take place too. For example, if it rains for a long time, it is more likely that you may get wet. However, *prediction* does not mean that an event is inevitable – other factors may influence the final outcome. If it rains for a long time and you go outside, you are more likely to get wet. However, if you stay indoors, the likelihood of getting wet will be low (unless the roof leaks). |
| Discovery sample | In a *genome-wide association study*, the *discovery sample* is the group of people that researchers study in order to look for DNA differences or *SNPs* that are associated with a particular trait or traits. Because each SNP only has a very small influence on individual development on its own, discovery samples need to be extremely large – ideally in the millions of people – in order to make it clear which SNPs have a real impact on development. GWAS with larger sample sizes typically yield more strongly predictive *polygenic scores.* This is because they have greater *statistical power.* |
| DNA and genes | Deoxyribonucleic acid (DNA) is an organic chemical that carries the genetic instructions for the development, functioning, growth and reproduction of all known organisms. It is present in the nucleus of each cell of the organism and is shaped in the form of a double helix, consisting of bases (Adenine (A), Thymine (T), Guanine (G), and Cytosine(C)) and phosphates. Both strands of the 'ladder' are held together by bonds between the bases: A binds with T and C with G. Each 'rung' is referred to as a base-pair.<br>Sequences of DNA form genes. They can vary in size from a few hundred DNA bases to more than two million bases. |
| Genetic propensity | This term is used describe people's likelihood of showing or developing a particular trait based on their genetic makeup. It is important to note that our genes do not determine outright who we are. Our DNA differences make certain outcomes more likely, but our experiences in the world also play an important role. For instance, a person with a low genetic propensity for reading enjoyment may still develop a love of reading if they are provided with positive and stimulating experiences of reading. Other terms include *genetic tendency* and *genetic predisposition*. |

| | |
|---|---|
| Genome | The genome is the complete set of genetic information of an organism. See also *DNA* |
| Genome-wide association studies | In a *genome-wide association study*, researchers collect genotype information and information about a particular trait or traits (e.g., intelligence test scores, years spent in education, height) from a very large group of people. First, researchers identify which individuals show the trait of interest. Second, researchers compare genotype data from people who show the trait against genotype data from people who do not show the trait. This is done to discover which *SNPs* are linked to showing the trait. Some SNPs may be more commonly linked to a certain trait than others SNPs are, suggesting that these SNPs may have a greater influence on trait expression. |
| Genotype | A genotype is a scoring of the type of genetic variant that exists at a particular location (i.e., a locus) in the genome. For example, the sequence of DNA bases at a specific location, such as CC, CT, TT. |
| Heritability | Heritability is the proportion of differences in a trait that can be explained by genetic differences between people in a population at a given time. Genetic and environmental influences exist together: Individual differences between people in all traits are heritable to some extent, but no trait is 100% heritable. That means, even if a trait is highly heritable, the environment is still important, because environmental factors support or limit the extent to which someone's genetic propensities for a trait can be developed. |
| Missing heritability gap | The heritability of a trait can be assessed by comparing individuals who share DNA and life experiences with individuals who share only DNA, only life experiences, or neither (see *Twin and adoption studies*). Since the DNA revolution, heritability can also be assessed by identifying DNA differences between people (e.g., *SNPs*) and generating *polygenic scores* to predict observable phenotypic differences between people. Genotyping studies have so far yielded much smaller heritability estimates than twin and adoption studies.<br>This is because genetic research is a relatively young and still-developing field. This discrepancy is known as the *missing heritability gap* and it is encountered in genetic research across the social sciences. However, new methods are constantly being developed, which are gradually narrowing the *missing heritability gap*. |
| Phenotype | Phenotype refers to a person's observable or expressed traits. Differences between people in these observable traits are the result of their genes as well as the environments they experience.<br>In GWAS, the trait for which SNP associations are tested is called 'target phenotype'. |

| | |
|---|---|
| Pleiotropy | Pleiotropy refers to the phenomenon of one single genetic variant affecting two or more distinct phenotypic traits. That means, genetic influences are then shared between or common to these phenotypes, not specific to just one. |
| Polygenic scores | If we have genotype information for a person or group of people – that is, if we know what *SNPs* they possess – we can create *polygenic scores* for them, to predict how likely they are to show certain traits (e.g., height, colouring, intelligence etc). Polygenic scores are created by adding together the effects of all of the SNPs a person possesses that are associated with a particular trait. The SNPs associated with phenotypic traits of interest are identified using *genomewide association studies*. |
| Polygenicity | Most traits that we can observe (like height, colouring, musical aptitude, intelligence and so on) are influenced, not by one SNP alone, but by many thousands of them across many different genes, working together. Each SNP only has a very small influence on development, but certain combinations of SNPs can cause measurable differences between individuals, such as differences in physical appearance, personality, or academic success. This phenomenon is termed *polygenicity* and traits like these are described as *polygenic.* |
| Singlenucleotide polymorphisms (SNPs) | A SNP (pronounced 'snip') is a base-pair difference in the genome, which occurs at one location of the genome (i.e., in a single nucleotide), independent of basepair differences in other locations, and presents in many (i.e., poly) different versions (i.e., morphisms). SNPs are the smallest genetic variants that can differ between people. The human genome includes about three billion base-pairs, and SNPs occur about once in every 300 base-pairs (i.e., there are about ten million SNPs within the human genome). |
| Socioeconomic status | Someone's socioeconomic status (SES) refers to the person's position in society in terms of their social standing or class in relation to others. It is often measured as a combination of educational, occupational and economic (e.g., income) criteria. |
| Statistical power | The *statistical power* of a study refers to the likelihood that the study will be able to identify the effect or effects of what that study is testing if, indeed, there truly is an effect. If a study has high statistical power, there is a stronger chance that its results will be reliable, and vice versa. When the effect/s being measured is/are very small, a very large sample of data is needed to determine whether the effect under investigation is really there. If a study uses too small a sample, there is a risk that the study will fail to capture the relationship between the phenomena studied and their outcome/s. A null finding like this resulting from a study with low statistical power is termed a false negative or a Type II error. |

| | |
|---|---|
| Twin and adoption study designs | The classic twin study design relies on comparing twin similarity on a given trait between identical (monozygotic) and fraternal (dizygotic) twins. Monozygotic twins share 100% of their genes, while dizygotic twins share on average 50% of the genetic variants that vary between humans; they resemble each other genetically as much as other biological siblings. If monozygotic twins are more similar to each other in a given trait than dizygotic twins, we can conclude that genetic factors influence people's differences in that trait.<br>In sophisticated statistical models, we can use these twin correlations from identical and fraternal twins to calculate how much of the variation in a trait in the population is explained by genetic factors and how much is due to environmental factors.<br><br>In adoption studies, the similarity between adoptees and their biological as well as adoptive families is compared. If children were adopted early in life, similarity with their biological parents and siblings is entirely due to genetic factors, while similarity with their adoptive family members is fully accounted for by shared environmental factors, because they are genetically unrelated. |

## Introduction

### *Genetic and environmental contributions to individual differences*

The human *genome* describes the complete set of genetic information (DNA) that an organism possesses. The DNA differences that children inherit from their parents contribute to their phenotypic differences in physiological and psychological traits throughout the life course. That is to say, all traits are *heritable* to a certain extent (Turkheimer, 2000; Polderman et al., 2015). Amongst these heritable traits, intelligence (i.e., our ability to reason, think, and learn) and social-emotional behaviour (i.e., how we cultivate relationships with others) have particularly far-reaching impacts for an array of life outcomes. These include how much time a person spends in full-time in education, how many qualifications they attain, what their employment prospects are, their socio-economic status, their access to civic and social life, and their mental and physical health and wellbeing (Deary, 2012; Duckworth & Gross, 2014; Goodman et al., 2015; Fawns-Ritchie et al., 2017). Individual differences in these domains begin to manifest during early years development and become magnified across the lifespan (von Stumm et al., 2019). Behaviour genetic research can offer valuable insights into how and why these differences emerge and develop.

Genetic influences do not act alone or independently of the environmental context in predicting individual differences in phenotypic development. Rather, DNA differences act in concert with our everyday environmental experiences throughout our lives (Turkheimer, 2000; Dick, 2011; Polderman et al, 2015; Plomin et al, 2016; Black, 2017; Cattan et al, 2021). No two people have precisely the same set of experiences and interactions with the environment, even if they are genetically identical (i.e., monozygotic twins). Experiences that siblings in a family share, like parents reading a bed-time story to all siblings, make up the shared environment and explain siblings' similarities in cognitive and social-emotional outcomes. By contrast, things experienced by one sibling but not by others (e.g., one sibling joining the local library, while the other sibling joins a sports club) make up the non-shared environment. Non-shared environmental experiences can explain why two siblings of a family might show less similar or more different traits and life outcomes.

Understanding how our genetic make-up and our environmental experiences shape us cognitively and socially will, in the long term, help us to identify individuals who may benefit from additional support and enable us to deliver truly effective support within the appropriate

timeframe. Promoting equity of life outcomes across society is a social justice issue and merits our urgent attention.

### *Heritability*

The term *heritability* denotes the proportion of individual differences – that is, of *variance* – in a given trait that can be explained by genetic differences between people (Plomin, 2018; Visscher et al., 2008). Intelligence is around 50% heritable across the lifespan, increasing from 20% in infancy, to 41% in childhood, to 66-80% in early to late adulthood (Haworth et al., 2010; Polderman, 2015; Plomin and Deary, 2015). That is to say, genetic makeup can explain, on average, 50% of the differences in intelligence that can be observed between individuals. Note that heritability *does not* explain what proportion of a given trait in a given individual is inherited, so genetic makeup *does not* explain 50% of a given person's overall intelligence. It is also important to note again that heritability does not mean immutability. Genetics makes us *prone* to develop certain traits but biology alone does not determine who we are and how we develop. Rather, as mentioned earlier, the expression of phenotypic traits is influenced by the interplay between genetics and the environment.

### *Early approaches to inferring heritability*

Earlier behaviour genetic research used twin and adoption studies to investigate the heritability of a range of traits including intelligence, educational attainment and achievement, and social-emotional behaviour (Knopik et al., 2017). Comparing trait development in identical (*monozygotic*) and non-identical (*dizygotic*) twins allows us to gauge to what extent individual differences in psychological traits can be attributed to genetic and environmental factors. When raised together, both identical and non-identical twins share a rearing environment (e.g., their family home). However, identical and non-identical twins differ in terms of their genetic similarity to one another. Identical or monozygotic twins shared 100% of their genetic makeup, while fraternal or dizygotic twins share only 50% of their segregating genes. If identical twins are more similar to one another in a given trait – for example height – than are dizygotic twins, then we can conclude that differences seen in the expression of that trait are influenced by genetics (Turkheimer, 2000; Moore & Shenk, 2017; Polderman et al., 2015).

Likewise, comparing trait resemblance across biological siblings and adopted siblings, who are not biologically related, can offer insights about the extent to which traits are transmitted from generation to generation through biological or experience-dependent mechanisms. Biological siblings share, on average, 50% of their segregating genes and their entire rearing environment, whilst adopted siblings share only the latter (assuming adoption takes place vary early in life) (Knopik et al., 2017). If both biological and adopted siblings show strong similarities in a given trait, then these similarities most likely stem from shared environmental factors. If biological and adopted siblings differ substantially in a given trait, then we can conclude that individual differences in these traits can be largely attributed to genetic differences between the siblings.

### *Candidate gene approach: Early attempts at identifying DNA-based predictors*

Candidate gene studies focused on small groups of genes that were hypothesised to be responsible (i.e., candidates) for the expression of a given trait. Candidate genes were typically identified by examining a limited number of pre-specified genes using case-control studies. These candidate genes were detected in individuals with a given trait (cases), who were compared with individuals (controls) who did not show that trait to test whether candidate gene frequencies were higher in cases than controls (Kwon & Goate, 2000). Candidate gene studies can be effective in identifying *alleles* (or sets of *alleles*) in genes associated with diseases. Alleles are the alternative variants of a gene, inherited from our parents, that code for a certain trait e.g., blue versus brown eyes. For example, two rare variants of the BRCA gene, BRCA1 and BRCA2, have been shown to substantially increase the risk of breast, ovarian, and other cancers in women who carry this gene (Chen & Parmigiani, 2007). A key limitation of candidate gene research is that it relies on researchers selecting single genes to investigate *a priori*, and does not, by design, consider the whole genome. Most phenotypic traits, including psychological traits like cognitive development, are complex or *polygenic*, meaning that they are influenced by many thousands of DNA variants across many genes, not just one or two genes. Thus, traditional candidate gene studies excluded many genetic variants that are likely to play a role in trait expression. As a result, findings from candidate genes have largely failed to replicate, and significant candidate gene effects on developmental differences are now, in hindsight, mostly understood to be false positives (e.g., Chabris et al., 2012, Culverhouse et al., 2018, Border et al., 2019; van de Weijer et al., 2022). For instance, a body of candidate gene

research has identified 18 individual genetic variants to be linked to the risk of major depression (Border et al., 2019). However, when these genetic variants were considered together, neither reliable links between participants' alleles and their diagnosis emerged, nor clear evidence of interactions between these alleles and participants' shared environmental experiences (Border et al., 2019).

### *DNA revolution: GWAS and Polygenicity*

The DNA revolution and the mapping of the human genome made it possible to examine the specific genetic differences that drive some of the heritability that is estimated in twin and adoption studies. This alternative to the candidate gene approach involves collecting genetic information from across the genome to aggregate DNA variants, or *single-nucleotide polymorphisms* (SNPs, pronounced 'snips'), that are associated with a given trait. This is the method utilised in the present project.

Genetic variance in complex traits, like cognitive and social-emotional development, is due to many thousands of single alleles or *SNPs* at positions across the human genome that are associated with a target trait. This phenomenon is termed *polygenicity*. SNP associations with a target trait have very small effect sizes, meaning that they can only be detected in studies termed *genome-wide association studies* (*GWAS*), which have very large discovery samples. In a GWAS, data from hundreds of thousands of individuals, who underwent genotyping and phenotypic testing, are pooled together. Testing for psychological phenotypes usually involves, for example, intelligence tests, surveys on psychological traits and behaviour, and clinical diagnoses. Researchers can then test whether there are associations between individual DNA variants across the genome and individual differences in traits between people (Uffelmann et al., 2021). Using the association data generated by GWAS (i.e., their summary statistics), person-specific *polygenic scores* can be computed in any sample that includes information on participants' genotype, for example from saliva or blood samples (Choi et al., 2020). To compute polygenic scores, all of the many small effects of each person's SNPs, which were found to be associated with the target phenotype in a GWAS, are aggregated and, typically, weighted by their strength of association with the target phenotype to create the polygenic score (Plomin & von Stumm, 2018).

Polygenic scores for certain traits can be used to help to identify individuals at risk of developing diseases, learning difficulties, and developmental delays in a range of domains

(e.g., physical, cognitive, and social development). However, as discussed previously with regard to heritability, a person's polygenic score does not determine outright whether they will show a given trait. Rather, a polygenic score is a measure of *genetic propensity*, indicating their likelihood to show that trait. Additionally, although genetic propensities are connected with environmental experiences, a polygenic score does not, in and of itself, capture to the role of environmental influence in trait expression (see Pingault et al., 2022).

### *Statistical power*

Because SNPs show very small effect sizes, it is necessary for GWAS to recruit very large discovery samples of participants in order to identify trait-associated SNPs. GWAS with larger sample sizes typically produce more strongly predictive polygenic scores, since they are statistically better *powered* to detect these small effect sizes (Wilding et al., 2023). Recent GWAS of educational attainment (measured by asking participants how many years they had spent in education) recruited discovery samples of 1.1 million (Lee et al., 2018) and >3 million participants (Okbay et al., 2022) and have yielded polygenic scores predicting 11% and 13% of variance in phenotypic educational success, respectively. By contrast, the largest GWAS of intelligence to date (Savage et al., 2018) recruited a discovery sample of ~270,000 and the resulting polygenic score predicted a weighted mean of 4.8% of variance in cognitive ability across independent samples in this GWAS (Oxley et al., 2023).

### *Missing heritability*

Polygenic scores at present predict a smaller portion of heritability than twin and adoption studies. The upper limit of variance that could be explained by polygenic score predictions is defined by the SNP heritability. This indexes the proportion of phenotypic variance explained by all measured SNPs in a GWAS. SNP heritability is, on average, lower than twin study estimates of heritability but higher than polygenic score prediction effects. The difference between twin heritability, SNP heritability, and polygenic score prediction in the proportion of variance explained is termed *the missing heritability gap*, and is seen in behaviour genetic research across the social and psychological sciences (Yang et al., 2011; Plomin & von Stumm, 2018). Given that larger sample sizes typically give rise to more predictive polygenic scores, it is possible that the missing heritability gap will narrow as further GWAS with increasingly larger sample sizes are undertaken (Yengo et al., 2022). The predictions from presently available

polygenic scores offer valuable insights into the genetic causes of individual differences, although their predictive powers are probably lower bounds and underestimate the prediction from future polygenic scores.

### *Gene-environment interplay*

The relationship between genetic and environmental factors influencing early years cognitive development is complex and driven by two key processes: *gene-environment correlation* (rGE) and gene-*environment interaction* (GxE). rGE describes how our individual genetic differences are linked to our individual environmental experiences. GxE explains how genetic differences can give rise to different outcomes in individuals who share their environmental experiences, and how non-shared environmental experiences can lead to different outcomes in individuals who share a common genetic propensity to a given trait.

Gene-environment correlation denotes the fact that individuals, including children, are not randomly assorted across environments. Instead, individuals select, shape, and create the environments that match their genotypes (Plomin et al., 1977). Our genetic makeup correlates with our environmental experiences in three different ways – passively, evocatively, and actively. Passive rGE describes how children are experiencers of environments created for them by their parents. The environments that parents create are shaped by their own genetic predispositions which were, in part, transmitted to their offspring. For example, parents who enjoy reading may pass this tendency on to their children via inherited DNA differences, and may also encourage the reading of stories at bedtime, which indirectly strengthens children's genetically predisposed literacy skills. Evocative rGE describes how individuals bring about certain environmental experiences by acting on or in their environment, influenced by their own genetic predispositions. For example, children who are genetically predisposed to enjoy reading may show great interest in reading and this might cause them to receive more support and encouragement from parents and teachers to further develop this tendency. Active rGE refers to the fact that our genetic makeup predisposes us to prefer and select certain environmental experiences. For example, children who are predisposed to enjoy reading may visit the library more often or spend their pocket money on books, while children who find reading less enjoyable may prefer to focus on other activities that interest them more.

Gene-environment interaction denotes cases where people with similar genetic predispositions show different traits and outcomes according to their environmental experiences, and where

similar environmental experiences may give rise to different traits and outcomes in people with different genetic predispositions. For instance, different children may benefit from reading support in school or at home to different extents depending on their genetic predisposition for reading skills and reading enjoyment (Plomin et al., 1977; Price & Jaffee, 2008; Anreiter et al., 2018). It is important to note that identifying a gene-environment interaction can only tell us about the *presence* of that interaction and not about the *directionality* of the interaction. That is, we cannot determine whether the developmental impact of environmental factors varies between individuals according to their genetic make-up, or whether the developmental impact of genetic make-up varies between individuals as a result of differing environmental experiences. There are three conceptual frameworks for understanding gene-environment interaction: the *diathesis-stress model*, the *bio-ecological model*, and the *model of differential susceptibility* (also see Figure 1).

### *The diathesis-stress model*

This framework takes the view that negative environmental experiences affect individuals differently based on their genetic propensity for specific developmental outcomes (Zuckerman, 1999; Dick, 2011; Sigelman & Rider, 2009). That is, children with a stronger predisposition towards cognitive developmental outcomes, for example, may be more resilient to environmental experiences that are either harmful or not conducive to cognitive development, like receiving little or insufficient cognitive stimulation in early life. By contrast, children with a weaker genetic propensity for cognitive development may be disproportionately vulnerable to such experiences. For instance, children with a weaker propensity for literacy or verbal reasoning may be more significantly disadvantaged if they have little exposure to reading practice in their early life, such as joint book reading (cf., Noble et al., 2019) or bedtime story telling, than children with a greater propensity for these traits. This difference may contribute to the emergence of delays or differences in written and oral communicative development.

### *The bio-ecological model*

This model proposes that development results from cyclical transactions between the genome and the environment (Bronfenbrenner & Morris, 2006; Pluess & Belsky, 2013). The extent to which a given trait is expressed is influenced by interactions between an individual's genetic propensity for that trait and their environmental experiences that might stimulate or suppress

that propensity (i.e., their interactions with their families, peers, nurseries, schools, and the wider world). Some environmental experiences may increase the likelihood that a trait will be expressed, whilst others may impede the expression of the same trait. For instance, a child with a strong genetic predisposition for literacy or verbal reasoning skills who is exposed to a stimulating reading environment (e.g., a child who is often read to or encouraged to read) may become proficient at reading but the same child's inclination to read might be suppressed if they experience little or no exposure to reading. In this case, enriched environments would maximise the expression of genetic differences, while scant environments would mask them (Figure 1b).

### *The model of differential susceptibility*

This framework posits that children with greater genetic susceptibility to environmental influences may be disproportionately affected in their development when exposed to either positive and negative environments (Greven et al., 2019). For example, reading books with parents frequently may greatly benefit language development in children who have inherited a greater sensitivity to the environment, while less joint book reading may badly impair language development in the same children (Figure 1c). By contrast, in children who are genetically less susceptible to the environment, language development may be relatively unaffected regardless of the degree to which reading is encouraged by adults in their environment. This notion is central to theories of environmental sensitivity, which propose that individuals differ in their sensitivity to both aversive as well as supportive environments (Belsky & Pluess, 2009).

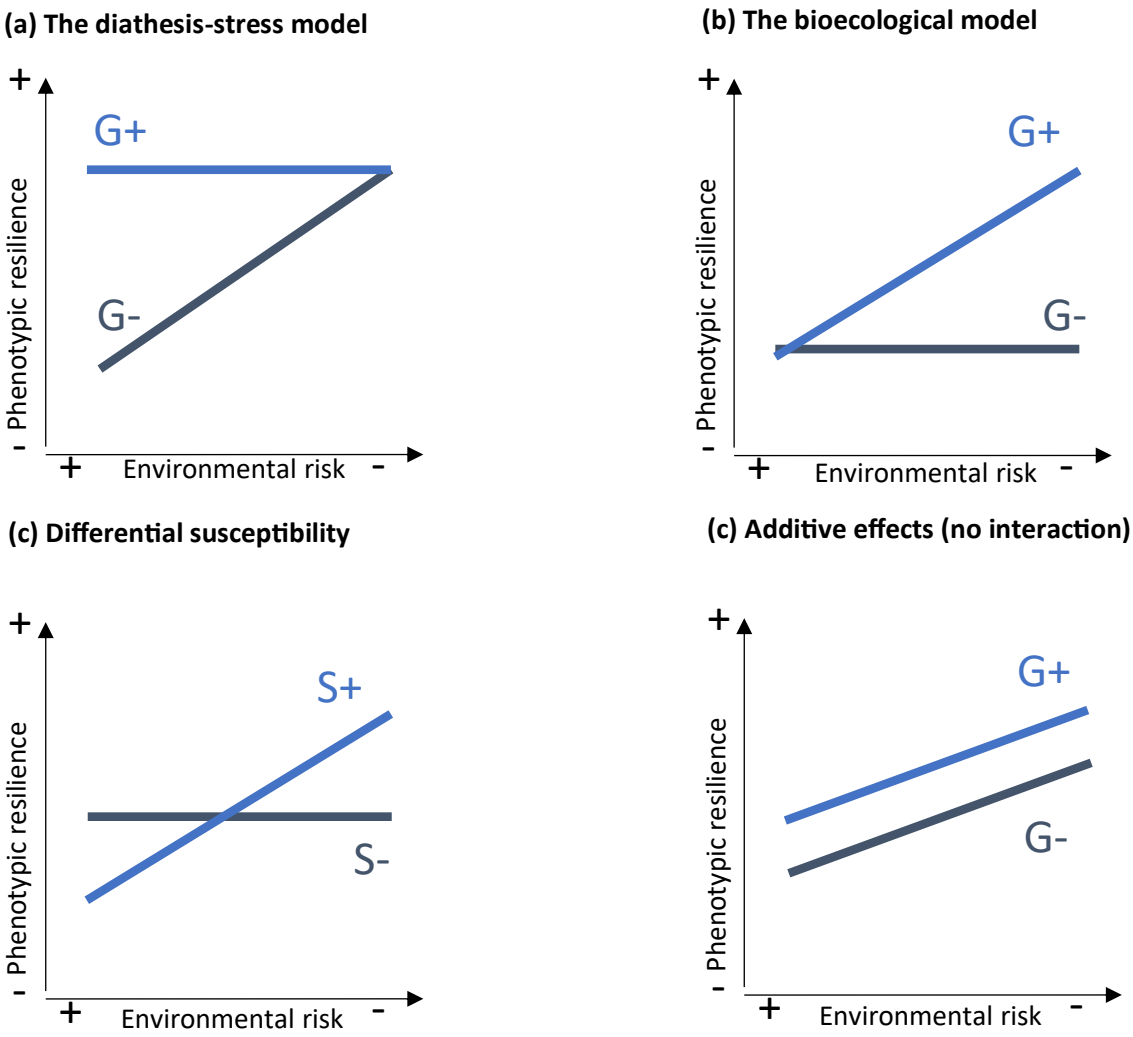



Figure 1. Models of the gene-environment interplay

Note. Panel (a) illustrates the diathesis-stress model, where negative environments have detrimental effects on the phenotypic development of individuals with greater genetic risk (G-) but not those with low genetic vulnerability (G+). Panel (b) shows the bioecological model, where environmental stress masks genetic differences, so that genetic differences become phenotypically expressed in positive but not in negative environments. Panel (c) demonstrates differential susceptibility, where negative and positive environments have disproportionately larger effects on individuals with greater sensitivity to environmental influences (S+) compared to those with low environmental sensitivity (S-). Panel (d) shows additive effects of genetic and environmental factors without gene-environment interaction. Adapted from von Stumm and Nancarrow (2023).

*Challenges in gene-environment interaction research*

Whilst it is now well-established that the interplay between genetic makeup and environmental experience is critical in shaping development, previous research has struggled to produce replicable, robust evidence of gene-environment interactions (e.g., Tucker-Drob & Bates, 2016). One possible reason for this is that previous studies may not have had the statistical power that is needed to detect gene-environment interactions. Such interactions typically have small effect sizes, meaning that very large participant samples (i.e., in the hundreds of thousands) are required to in order to detect significant effects and reject the null hypothesis (Cohen, 1992a, 1992b). Another possible reason is that, while there are many environmental factors that are known to influence development, previous genetically sensitive research has incorporated only a narrow range of these factors, focusing on only one or two variables (Schmitz & Conley, 2017). Thus, may environmental factors have yet to be considered in GxE studies. Additionally, previous research on GxE effects did not systematically incorporate polygenic scores, but instead relied on candidate gene and twin and family study designs.

The present project aimed to address these lacunae by testing for interactions between genetic and environmental factors known to affect cognitive and social-emotional development, using polygenic scores derived from the summary statistics from a range of recent large GWAS, including those for neuroticism, ADHD, and educational success (Luciano et al., 2018; Demontis et al., 2019; Lee et al., 2018), and data pertaining to a wide range of environmental variables in a longitudinal, UK population representative cohort study of children born between 1994 and 1996 in England and Wales.

We used Lee et al.'s (2018) polygenic score for educational success as a proxy for genetic propensity for cognitive ability. While it was designed to predict educational success rather than intelligence *per se*, these polygenic scores actually predict a higher proportion of peoples' differences in intelligence than any polygenic scores for intelligence to date. This is because the GWAS that produced these polygenic scores (Lee et al., 2018) used a much larger discovery sample of ~1.1million than the largest GWAS of intelligence to date (Savage et al., 2018), which used a discovery sample of only ~270,000. Measuring educational success is relatively straightforward, involving only a single self-reported item (i.e., 'How many years did you spend in education?') compared with measuring intelligence, which requires complex and time-consuming IQ testing. Larger discovery samples are therefore more feasible for the former. We used Luciano et al.'s polygenic score for neuroticism (N~329,821) and Demontis et al.'s

polygenic score for ADHD (N~55,374) as measures of genetic propensity for social-emotional traits. Gidziela et al.'s (2021) meta-analysis assessing the predictive validity of 15 distinct polygenic scores for social-emotional traits established that these polygenic scores, amongst others, are positively associated with parent and teacher ratings of children's social emotional behaviour (see *Methodology: Sample* below).

Identifying gene-environment interactions in early years cognitive development could offer important insights into how we can best support children at risk of developmental difficulties and delays. Gathering knowledge about how inherited DNA differences and environmental experiences jointly influence the expression of cognitive and social-emotional traits during early life development will provide an evidence base for researchers and practitioners to work collaboratively to understand whether and how personalising learning could improve outcomes for some or all children. In the long term this could enable the design and implementation of truly effective therapies and interventions, within the most appropriate timeframe. Understanding which outcomes are or are not associated with which gene-environment interactions could, in this way, help us to identify at-risk individuals early in life, and to devise personalised programmes to ameliorate cognitive and social-emotional development (e.g., literacy and verbal reasoning skills, behavioural tendencies and well-being). The fact that cognitive and social-emotional development predict a wide range of life outcomes makes supporting individuals at risk for unfavourable development in these domains a clear priority for societies that aspire to fairness. Understanding how and why individual differences in cognitive and social-emotional development manifest is an essential step in working towards achieving greater equity in life outcomes for all members of society.

*Aims and Research Questions*

The present project set out to conduct exploratory empirical research to identify gene-environment interactions (GxE effects) in children's early life cognitive and social-emotional development. In doing so, we also built on previous research by elucidating to what extent cognitive and social-emotional development can be predicted by rGE and by environmental experiences directly. However, our specific aim for this project was to test for GxE effects during early life cognitive and social-emotional development by examining the combined predictive power of polygenic scores and environmental factors known to influence development. We sought to address the following research questions:

i. Can GxE effects in children's early life cognitive and social-emotional development be identified?

ii. If GxE effects can be identified, to what extent do they explain children's differences in early life cognitive and social-emotional development, independent of additive genetic and environmental effects?

iii. Which environmental variables, if any, interact most strongly with polygenic scores in the prediction of children's differences in early life cognitive and social-emotional development?

Because previous research has not produced reliably replicable GxE effects, our analyses rested on the following assumptions rather than testing specific hypotheses:

i) That our samples had adequate statistical power to produce reliable estimates of GxE effects across all environmental variables included in our analyses

ii) That confounding would be reduced by testing simultaneously for multiple interactions between polygenic scores and a wide range of environmental variables, which assessed different levels and domains of environmental influences

iii) That any identified GxE effects should also be identifiable (i.e., replicable) using alternative modelling approaches and additional samples.

## Methodology

### *Sample: The Twins Early Development Study*

We capitalised on data from the Twins Early Development Study (TEDS), the world's largest genetically sensitive study of development, that initially included 13,759 families with twins born between 1994 and 1996 in England and Wales. When twins reached 18 months of age, families provided data about demographics, pregnancy, childbirth, and zygosity (i.e., whether twins are identical or fraternal). During subsequent assessments when the twins were aged 2, 3, and 4 years, data were collected on cognitive and social-emotional development and environmental factors, including parenting, the family home environment, and early life experiences. At ages 7, 9, and 12, the twins' teachers provided information about their social-emotional behaviour in school. DNA samples were collected throughout the twins' childhood, which are the basis for generating polygenic scores that indicate the twins' genetic propensities for later educational attainment and various cognitive and social-emotional developmental outcomes.

Table 1 provides an overview of the representativeness and data availability in TEDS (adapted from Rimfeld et al., 2019). Of the families who contributed to the first assessment wave, 83.5% (i.e., 10,336 families) also provided data between the ages 2 and 4 years, including 6,287 twin pairs for whom genotype data are available. After stringent quality control, the sample size available for the proposed analyses is 10,117, including 6,797 unrelated individuals and 3,320 dizygotic co-twins (i.e., non-identical twins). Although TEDS has suffered some attrition over the years, the analysis sample for this project differs only slightly in demographic characteristics from national estimates for all UK parents with children born in the 1990s. For example, in the 1990s, 35% of mothers and 47% of fathers had achieved A-levels or higher educational qualifications (DfEE, 2000), compared to 40% and 47%, respectively, in the subsample from TEDS used in our analyses. Similarly, 50% of mothers and 91% of fathers in the 1990s are estimated to have been in full-time employment while their youngest child was below the age of 2 years (ONS, 2017; Rimfeld et al., 2019). In our analysis sample, the corresponding estimates are 45% and 93%.

An extremely conservative preliminary power analysis for our proposed investigations (conducted in G*Power; Faul et al., 2019) indicated that this sample of 10,117 exceeded the minimum sample required to detect GxE effects at 80% power (N=8,358), assuming. This estimate was based on the assumption that environmental variables would exert no direct causal effect on outcomes, and that 100 gene-environment interaction terms would exert very

small effects (1%). A more realistic yet still conservative estimate, assuming small direct causal effects of environmental variables (5%) and very small GxE effects (1%) indicated a smaller participant sample (N=3,755). Thus, while a small proportion of the participant sample have partially incomplete data on our 39 environmental variables of interest, the TEDS database was deemed to constitute a more than sufficient sample for our proposed analyses.

Table1. Sample properties and characteristics in TEDS

| **Returned families** | | **% Response** | **% White** | **% Mother A-levels +** | **% Father A-levels +** | **% Mother employed** | **% Father employed** | **% Girls** |
|---|---|---|---|---|---|---|---|---|
| **Full sample** | | | | | | | | |
| First contact | 13759 | 84.4 | 91.7 | 35.5 | 44.8 | 43.1 | 91.6 | 50.0 |
| Age 2 to 4 | 10336 | 70.5 | 92.9 | 37.8 | 46.3 | 43.2 | 92.2 | 50.7 |
| **Genotyped sample** | | | | | | | | |
| First contact | 7019 | 99.9 | 99.9 | 39.6 | 46.9 | 45.6 | 92.9 | 51.9 |
| Age 2 to 4 | 6287 | 90.1 | 99.9 | 40.1 | 47.1 | 45.4 | 93.2 | 52.0 |

*Note.* Numbers adapted from Rimfeld et al. (2019).

### *Study materials and data collection: Children's development and genetic propensities*

In TEDS, several measures for cognitive ability were available longitudinally at the ages 2, 3 and 4 years, using established, standardised scales that allowed us to analyse children's early development. In addition, data on anxiety and behaviour problems were collected as well as on physiology, including anthropometric data that were not relevant to our analyses.

*Cognitive development.* Nonverbal cognitive performance was assessed using age-appropriate versions of the Parent Report of Children's Abilities (PARCA). This is an hour-long test consisting of three parent-administered tasks: a "find the pair" task, a drawing task, and a matching task. Items in these tasks assessed number, shape, size, conceptual grouping, and orientation skills. Some items were newly created for administration in TEDS, while others were adapted from previously well-validated measures, such as the McCarthy Scales of Children's Abilities (McCarthy, 1972) or the Bayley Scales of Infant Development (Bayley, 1993). The parent-administered PARCA component was supplemented by eight parent-report items anchored on concrete behaviours and requiring simple 'yes' or 'no' answers. Some of these items were newly created; others were adapted from the Minnesota Child Development Inventory (MCDI; Ireton & Thwing, 1974) and the Ages and Stages Questionnaires (Bricker,

Squires, & Mounts, 1995). Verbal ability, including vocabulary and grammar, was assessed by parent reports for the CDI-III, an extension of the short form of the MacArthur Communicative Development Inventories: Words and Sentences (Fenson et al., 2000). The PARCA has undergone testing and is established to be a valid and reliable measure of children's cognitive abilities (Bayley, 1993; Blaggan et al., 2014; d'Apice, Latham, & von Stumm, 2019; Martin et al., 2013; McCarthy, 1972; Oliver et al., 2002; Saudino et al., 1998). The United Kingdom's National Institute for Clinical Excellence (NICE, 2017) uses a revised version (PARCA-R) in their developmental assessment guidelines. For our analyses, at each age, scores were standardised and summed to form z-scores of cognitive ability.

*Social-emotional development.* Information on children's behaviour problems was captured by the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). The SDQ is a brief screening questionnaire asking about attributes of children's behaviour divided into four scales: Emotional symptoms, peer relationship problems, hyperactivity/inattention, and conduct problems. We used these scales to assess teacher ratings of children's behaviour problems at ages 7, 9 and 12 years. Teacher ratings of behaviour problems were obtained via mail and included a total of 20 items at each age, that is, five items for each of the four scales. Teachers assessed, for example, if the child is "Restless, overactive, cannot stay still for long", is "Generally liked by other children" or is "Often argumentative with adults". The items were rated on a three-point Likert scale (certainly true; sometimes true; not true), wherein higher scale scores always indicated more challenging behaviour.

*Polygenic scores.* In TEDS, the twins' DNA was extracted from saliva and buccal cheek swab samples, collected in two waves, 5 years apart. Raw data from 10,346 individuals on over 600,000 SNPs underwent quality control and genome-wide information was imputed for subsequent computation of polygenic scores. In the analyses in this project, we used polygenic scores for educational attainment (measured as years spent in education; Lee et al., 2018), attention deficit hyperactivity disorder (ADHD; Demontis et al., 2019) and the personality trait neuroticism (Luciano et al., 2018). Because the genetic information encoded in the DNA (i.e., one's inherited DNA differences or SNPs) remains unaltered across the lifespan, the validity of polygenic scores is unaffected by the age at which DNA was collected. As discussed earlier, the polygenic scores that we used for our analyses have been found elsewhere to predict children's educational attainment and achievement (Wilding et al., 2023), cognitive ability (Allegrini et al., 2019), and childhood behaviour problems (Gidziela et al., 2022).

### *Measuring the environment*

To assess the environment, TEDS collected data on a number environmental factors that are important for developmental outcomes from birth up to early childhood, including child- and parent-reports of their experiences of family-, childcare-, and neighbourhood-level environments. Survey data on environmental factors was then linked with data from 'geo-coding', which captures sociodemographic data, such as deprivation, type of housing and households, air quality, health factors, life expectancy, crime statistics, and proximity to green areas, all based on UK postcodes.

To adequately address diversity in the environmental factors experienced by different individuals, we adopted an 'atheoretical' approach to model environments: We included as many environmental factors as were available in the data in our analyses to explore which factors predicted cognitive development. Just as we assumed for individual DNA variants or SNPs, we assumed that each environmental factor would only have a very small effect and, therefore, that these effects would need to be added together to explain any substantial proportion of variation in the trait of interest. For our analyses of gene-environment interactions in children's cognitive development, 39 environmental factors were available that broadly fell into four categories, including the home environment, family adversity, the neighbourhood environment, and environmental pollution variables (also see Figure 2). We summarise the measures included in each category measures in the following sections.

*Home environment.* Information on children's home environment includes parental reports on breastfeeding (number of days for which the child was breastfed) as well as the child's overall eating and sleeping behaviour. Parents reported the number of books, puzzles and children's tapes, records, and CDs available in the household and indicated how often the child engaged in various types of games (e.g., messy, physical, music, and board games). Information on parents' verbal interaction with the child (e.g., correcting pronunciation or sentence structure), as well as on interactions such as singing nursery rhymes and joint book reading were included. In addition, parents' provided ratings on their views on discipline, using the Parent Feelings Questionnaire (Deater-Deckard, 2000), and on chaos in the family home, using the CHAOS scale (Matheny et al., 1995). Lastly, the overall family socioeconomic status (SES) was assessed using information on mother's and father's educational qualification levels (1 = no qualifications to 8 = postgraduate qualifications), their employment status, and mother's age when they gave birth to their first child. These variables were included in the analyses both as individual variables, and in the form of a composite score.

*Family adversity*. Indicators of adversity included medical risk factors both for the mother, relating to pregnancy and birth (e.g., number of weeks gestation at birth, number of days prescribed bed rest during pregnancy, number cigarettes smoked per day and units of alcohol consumed per week during pregnancy), and for the twins themselves (e.g., birth weight, number of days in neonatal intensive). In addition, maternal postnatal depression was recorded and previous experiences of major life events were assessed. Life events included, for example, whether a family member has ever had a serious illness, whether the respondent's job or marital status changed, and new children living at home.

*Neighbourhood environment*. To assess the neighbourhood environment, several aspects were taken into account, including the overcrowding factor of the area (i.e., average household size divided by average number of rooms), occupancy ratings (i.e., proportion of households that have fewer rooms/bedrooms than inhabitants), the proportions of people who were living in social housing, who had no qualifications, who were in full- and part-time employment, and who were lone parents.

*Pollution variables.* To estimate air pollution in the area where children grew up, we included data on benzene, sulphur dioxide (SO2), nitrogen dioxide (NO2), nitrogen oxides (NOX) and particulate matter at less than 10 micrometres in diameter (PM10) and less than 2.5 micrometres in diameter (PM2.5). Benzene is used widely in major industrial manufacturing process, including plastics, foams, dyes, detergents, solvents, and insecticides. Other major sources of benzene include vehicle exhaust and the evaporation and manufacturing of petrol. Thus, high levels of benzene pollution are seen around industrial areas and roads. The majority of sulphur dioxide in the environment comes from burning of coal and oil at large industrial plants. Vehicle exhaust, domestic boilers and fires also contribute to increased levels of sulphur dioxide. Nitrogen dioxides are a group of gases that form mainly during the combustion of fossil fuels, mostly coming from vehicle exhaust and energy and manufacturing industries. Particulate matter (PM10 and PM2.5 here) covers all material in the air that is not a gas and is measured by the size of the matter. Around 15% of particulate matter in the UK comes from natural sources such as pollen and sea spray. Particulate matter can travel large distances, resulting in a third of particulate matter in the UK coming from other European countries. Half of the particulate matter in the UK comes from domestic wood burning and tyre and brake wear from vehicles. The modelled background air pollution data at 1x1 km resolution was obtained from the UK AIR website (https://uk-air.defra.gov.uk/data/pcm-data). These data have been routinely collected every year (since 2001/2002 depending on which pollutant) to provide policy support to Defra,

for reporting on air pollution levels to Europe, and are made publicly available for other research purposes such as epidemiology and public health research. More details on the modelling of the air pollution measures can be found online (Technical Report on assessments of air quality in the UK).

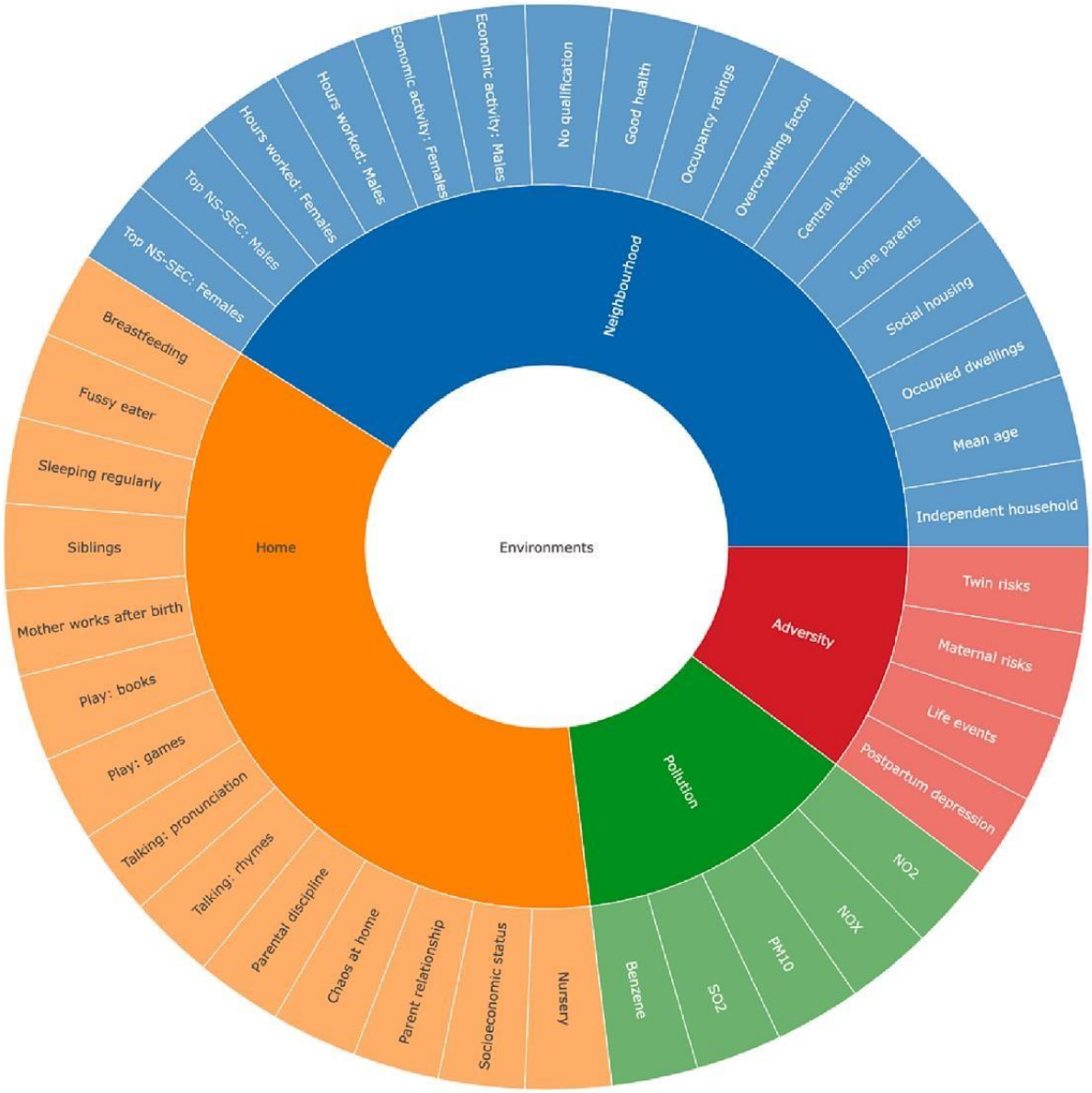



Figure 2. Environmental measures in TEDS, grouped by environmental domain for analyses in GxE in cognitive development. Adapted from von Stumm et al. (2023).

For our analyses on gene-environment interactions in children's development of behaviour problems (Plomin et al., 2022), we took a different approach compared to modelling interactions between the polygenic scores and the above outlined individual environmental factors. Instead, we selected eight parent-reported environmental variables that are related to behaviour problems in childhood from several hundred available in the TEDS data using two criteria. Firstly, we only included environmental factors that were at least moderately associated with children's behaviour problems and, secondly, we excluded highly correlated environmental measures. We began with two environmental risk measures at age 3 and at age 4: A family-general measure based on variables that are the same for co-twins (as a measure of shared environmental factors) and a twin-specific measure based on variables that differentiated between co-twins (as a measure of non-shared environmental factors). The family-general environmental risk composite included five standardised scores: Family SES, prenatal and perinatal medical risk, household chaos (the Confusion, Hubbub and Order Scale; Matheny et al., 1995), maternal postnatal depression the Edinburgh Postnatal Depression Scale (Cox et al., 1987), and life events such as changes to marital status, new siblings, mother's pregnancy, job changes and serious illness/accident). The twin-specific environmental risk measure included the same five variables as the family-general environmental risk measure, along with 17 additional variables detailing twin medical risk factors, parental discipline, and a parental feelings (Deater-Deckard et al., 1998). At each age we also included a separate item about smacking and shouting because of its predictiveness (Gidziela et al., 2022).

Because the environmental measures that we used here were correlated with one another across age 3 and 4, we grouped variables into two categories. The first category comprised general environmental risk factor and included both twin-specific and family-general environmental risk composites at ages 3 and 4, as well as SES. The second category comprised discipline and included the parental discipline composites and smacking/shouting variables at ages 3 and 4. The environmental risk and discipline categories correlated moderately. We used both categories as environmental predictors in the analyses.

### *Analytical approach*

To address our research questions, we systematically tested for GxE effects in TEDS. The main analyses used regression analyses to model interactions between environmental variables and polygenic scores for years spent in education (Lee et al., 2018), as a proxy for genetic

propensities for cognitive development, and polygenic scores for child ADHD (Demontis et al., 2019) and neuroticism (Luciano et al., 2018), as proxies for social-emotional development.

Specifically, interaction terms were usually first modelled at the level of single variables, the smallest available unit of measurement, and then tested jointly in the prediction of children's development. For predicting cognitive development, each environmental variable was included as a separate variable in our analyses, even in cases where they are known to be intercorrelated. For example, father's and mother's occupation, which are quantified using the census coding Standard Occupational Classification 2000; ONS, 2000) and correlate about 0.40, were initially each modelled with their own interaction terms, rather than including a family-level occupation variable that summarises multiple components (i.e., both mother's and father's occupation). This was to ensure that our findings provided the richest information possible about the specific mechanisms that are relevant in potentially very fine-grained effects of gene-environment interaction.

Although the proposed analytical approach relied on established statistical tools, no prior investigation has set out to simultaneously estimate a vast number of GxE effects. Thus, our approach was complemented by further statistical analyses, including (a) modelling GxE terms from composites for the environmental variable, and (b) investigating effects of gene-environmental correlation, which are highly relevant for finding GxE, by testing cross-lagged twin models that allowed us to infer causal, directional effects between variables.

For our analyses, we used polygenic scores to operationalise individuals' genetic propensities whereas previous studies have used candidate gene designs or indirect estimates of genetic influences from twin and family studies. Using polygenic scores is a decisive advantage of our approach with regards to granularity. The increased use of GWAS in recent years has shown that all complex traits, including cognitive and social-emotional development and educational outcomes, exhibit strong polygenicity. That is, the heritability of complex traits is driven not by a few individual genes, but by a large number of the smallest genetic variants (SNPs) across many genes, whose effects must be cumulated to explain individual variance (see Wilding et al., 2023 for an overview of DNA-based prediction of education-related traits).

Previous studies on GxE effects were also limited due to including only one or two environmental measures to capture children's rearing environment and any environmental experiences they encounter during their development. This is problematic since the effects of environments on development are likely to be at least as complex and manifold as that of

genetic information. Yet approaches to measuring and modelling the environment have not changed much for several decades, especially compared to the dramatic advances that occurred for the operationalisation of individuals' genetic propensities. Also, while inherited DNA sequences remain unaltered across age and have comparable units of measurement (i.e., SNPs), the environment that people experience changes in every moment and cannot be quantified in directly comparable units. Environmental factors are much more diverse and their influences are harder to quantify since they depend on how experiences are perceived subjectively by the individual (von Stumm & d'Apice, 2022). In our analyses, we address this challenge by including broad indicators of the environment as described above, which have been measured longitudinally and through ratings from multiple respondents where possible.

## Summary of findings

Using the approach outlined above, we have conducted several studies over the course of this project to investigate GxE effects in children's development and to answer our research questions. We have disseminated our findings in research papers that have been published in scientific journals, are pre-printed and submitted for review, or are in preparation. We list all research papers connected to this project above (see section Overview of scientific publications). Here, we briefly summarise the results and discuss their implications. For more specific information on the measures and statistical methods used in each study, the results reported in detail as well as an in-depth discussion of these findings, please refer to the published, open-access papers or preprints, respectively.

### *Do GxE effects exist in the prediction of children's early life cognitive development?*

von Stumm et al. (2023) analysed the relationship between the above-described range of 39 early life experiences, polygenic scores indexing children's genetic propensities for cognitive development, and composite measures indexing their phenotypic cognitive development, as tested at the ages 2, 3 and 4 years. We explored the main effects of these 39 environmental experiences and of polygenic scores for years spent in education (Lee et al., 2018) on cognitive development individually, as well as testing for potential interactions between these polygenic scores and each of these 39 environmental experiences. Specifically, we tested

whether GxE effects existed in the prediction of cognitive development in early life and if so, to what extent these GxE effects could account for children's differences in cognitive development that were not accounted for by direct effects of genetic or environmental factors alone. We also investigated which environmental variables were particularly relevant for GxE effects i.e., which of these variables explained most variance in children's cognitive development.

Our analyses revealed that all environmental measures together explained about 20% of variance in children's early cognitive development, with the strongest effects being observed for talking and playing with the child, the degree of family home chaos, and parental discipline. By itself, the educational attainment polygenic score explained 0.5% of the variance in cognitive development, but it no longer emerged as an independent predictor when added to a model that also included the environmental measures (Figure 3), suggesting that gene-environment correlations fully explained the independent prediction from polygenic scores. None of the tested GxE interaction effects for any environmental factor reached significance after controlling for multiple testing. Adding all interaction effects into the model with all environmental measures and the polygenic scores did not significantly improve the prediction of children's cognitive development beyond the direct effects of polygenic scores and environments. In short, we found no evidence for gene-environment interactions in the prediction of cognitive development from age 2 to 4 years. Yet, we found that the 'environmental' measures are substantially correlated with the polygenic score (–0.23 to 0.37), indicating that they are not truly external to the individual. Finding such strong evidence for gene-environment correlations (rGE) suggests that children's genetic propensities for cognitive development influence which environmental experiences they encounter and how they actively select, modify, and create their environments.

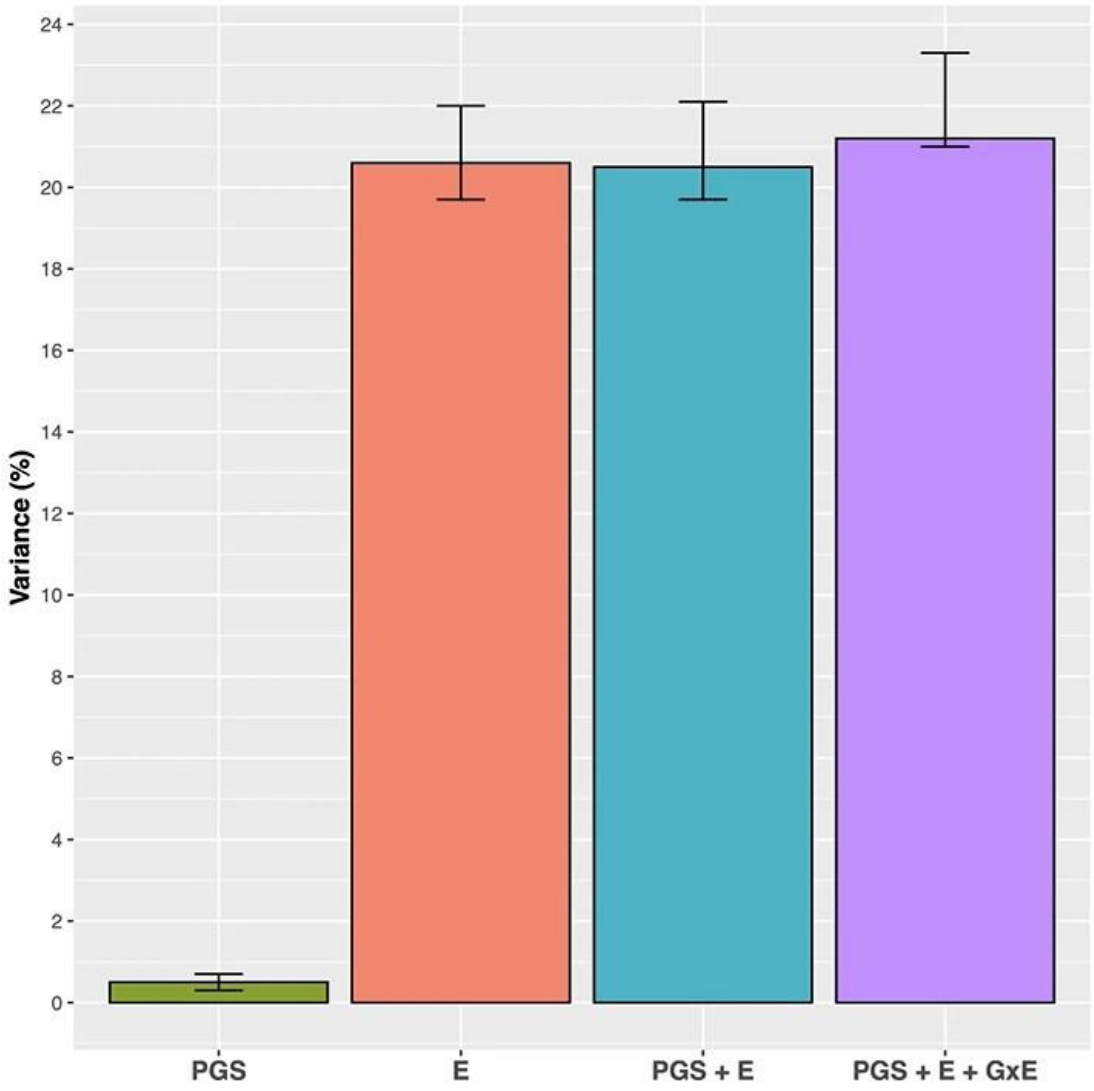



Figure 3. PGS and environment predictions of early cognitive development

Note. The bars show variance explained (adjusted $R^2$ in early cognitive development) by the polygenic score (PGS; first bar from left), all 39 environmental variables (E; second bar), PGS and E measures together (third bar), and combining all direct and indirect interaction effects from PGS and E (PGS + E + GxE: fourth bar). Adapted from von Stumm et al. (2023).

### *What drives the interplay between children's cognitive development and the environment they experience?*

While we did not identify robust GxE effects in children's development, we gained valuable insights into the mechanisms underlying the developmental association of environmental stimuli

and children's cognitive development. As putatively environmental factors are often assumed to exert causal influences on children's development, we tested the potentially causal relationship between children's cognitive development between the ages 3 and 4 years and the cognitive stimulation they receive from their parents (e.g., activities such as talking and rhyming, joint book reading and playing games; Oginni, Starr, & von Stumm, 2024). Receiving more cognitive stimulation is associated with more favourable cognitive development but it is unclear whether this association is due to a common origin (i.e., shared genetic or environmental factors throughout development) or due to causal directional paths and, if so, in which direction. Here, we aimed to clarify whether cognitive stimulation causes cognitive development or vice versa, and to shed light on potential common genetic and environmental influences.

Our analyses investigated prospective causal effects in both directions between cognitive stimulation and children's cognitive development (Figure 6). When controlling for correlations at each assessment and the developmental stability of cognitive ability and stimulation over time, we found consistent, bidirectional effects from cognitive development at age 3 to cognitive stimulation at age 4, and from cognitive stimulation at age 3 to cognitive development at age 4. However, these cross-lagged longitudinal effects were largely explained by underlying common genetic and shared environmental factors, rather than reflecting causal mechanisms. This indicates that contemporaneous mechanisms drive the causal influences between cognitive ability and stimulation: Children benefit from experiencing cognitively stimulating environments, and at the same time their cognitive development evokes the cognitive stimulation that their parents provide. Yet, interventions aiming to improve the stimulation children receive may only have limited purchase in directly influencing their cognitive development. These findings challenge long-held assumptions that children are simply passive recipients of the environments created for them by the adults around them (i.e., *passive rGE*). Rather, these analyses provided evidence that children exercise agency in co-creating cognitively stimulating environments tailored to their own genetic propensity for cognitive development, along with their parents and caregivers (*evocative rGE*).

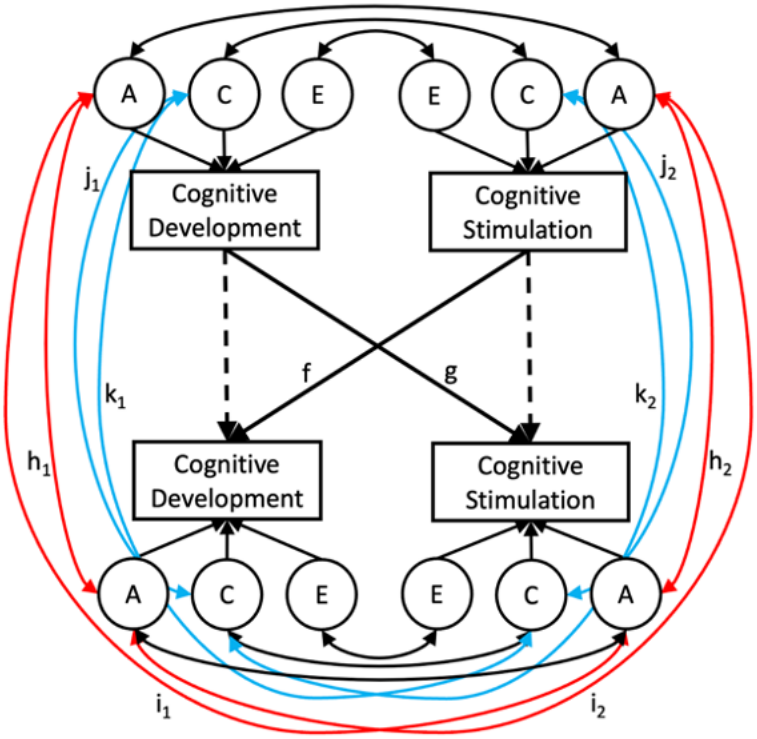



Figure 6. Cross-lagged model using twin data on the association between cognitive development and cognitive stimulation.

Note. The cross-lagged effects *f* and *g* are controlled for stability over time (dashed paths) and cross-sectional correlations (double-headed arrows) due to common genetic (A), shared environmental (C) and non-shared environmental (E) factors. The model additionally accounts for common genetic effects (red paths) and common shared environmental effects (blue paths) over time. Adapted from Oginni, Starr, & von Stumm (2024).

Our modelling approach allowed us to disentangle potentially confounding effects of common causes for both measures from true directional, causal influences. This represents a clear advantage over phenotypic, observational studies on children's development, which cannot account for mechanisms of genetic confounding. Estimating the proportions of causal influences and common causes in the overall relationship between children's cognitive stimulation and their cognitive development allows realistic estimates for the magnitude of the potential impact of interventions that seek to improve children's cognitive development. Our results confirmed that putatively environmental traits, such as the cognitive stimulation that children receive are genetically influenced. This finding is likely a result of gene-environment correlations, meaning

that, due to their genotype, parents may create rearing environments for their children that also correlate with said children's offspring genotypes, which in turn affects the children's phenotypic development. Our findings contribute to a growing body of evidence showing that children are not passive recipients of environmental inputs but select, modify, and create experiences that are correlated with their own genetic proclivities (Bronfenbrenner & Morris, 2006; Wertz et al., 2019).

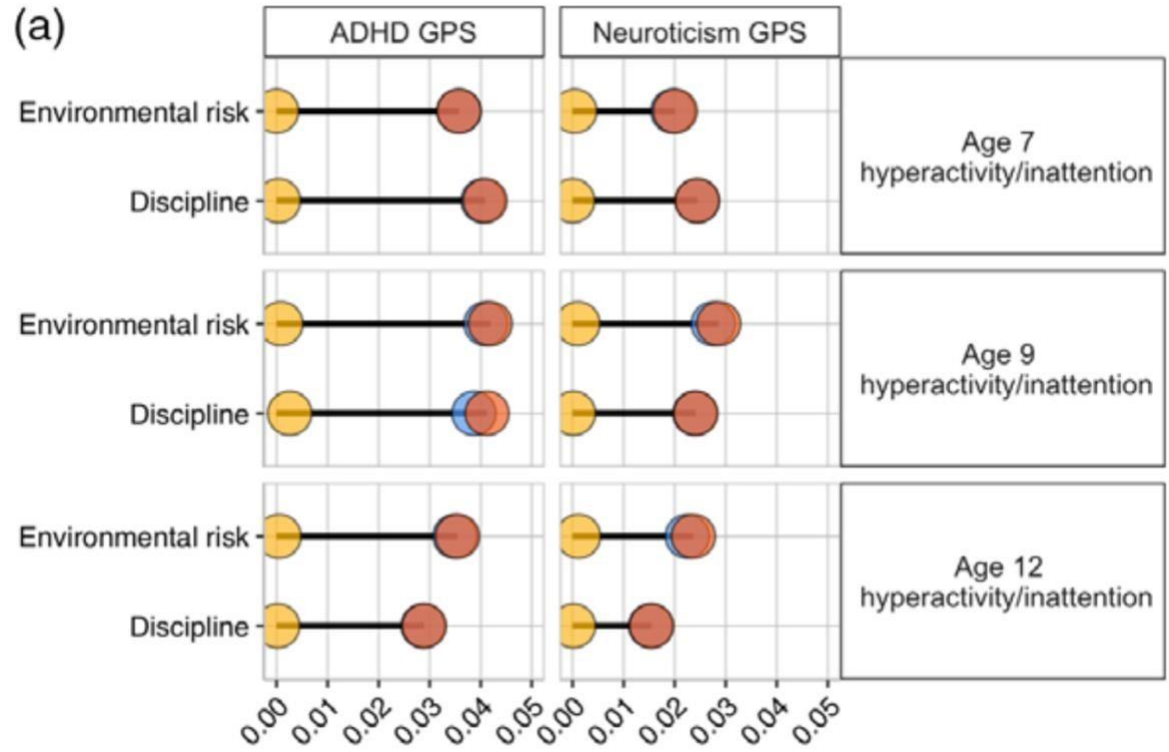
### *Are GxE effects relevant in predicting children's social-emotional development?*

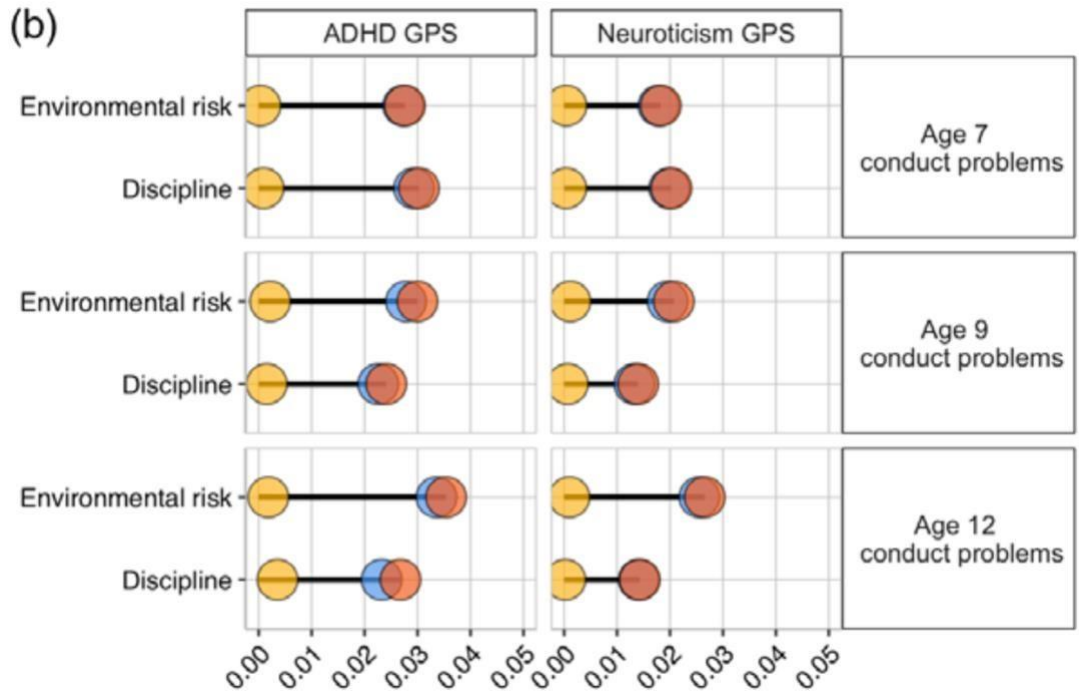
Plomin et al. (2022) investigated effects of gene-environment interactions for children's social-emotional development, in terms of hyperactivity/inattention, conduct problems, emotional symptoms and peer relationship problems (assessed using the Strengths and Difficulties Questionnaire; SDQ). These behaviour problems are known to be related to the environmental risks that children experience early in life, such as adverse family home conditions. However, it is unclear so far if children's genetic propensities for psychopathological disorders moderate this relationship.

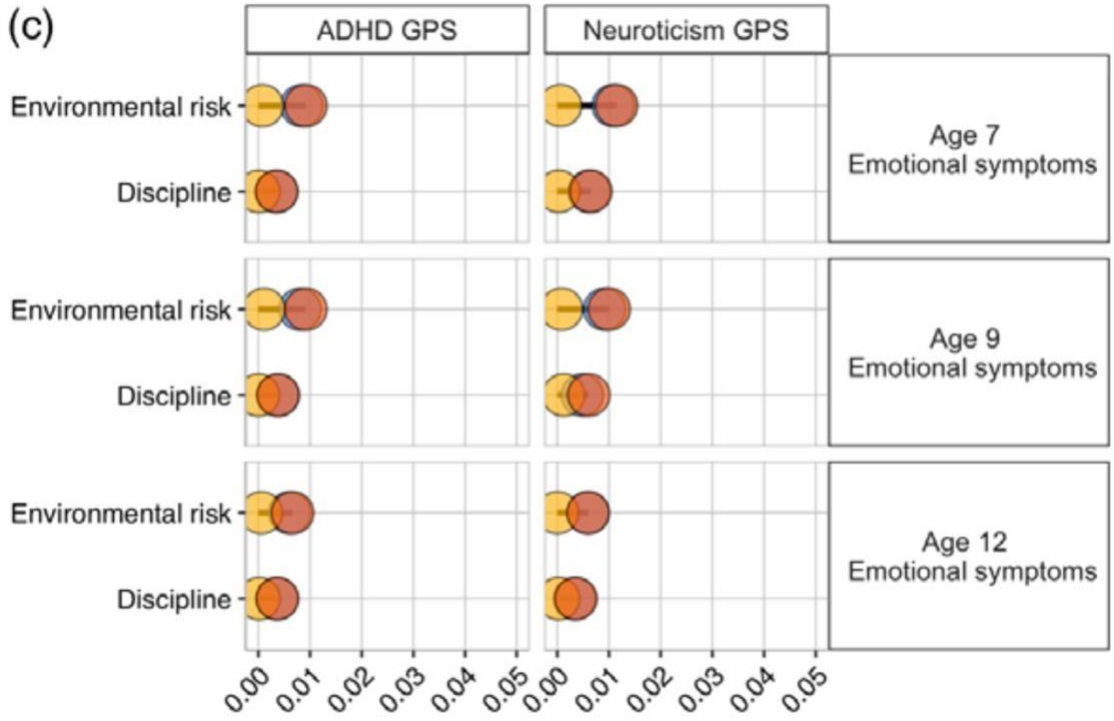
Hence, we tested to what extent children's problem behaviour could be predicted by environmental risk (E), polygenic scores for ADHD and neuroticism (G) and their interactions (GxE). We also investigated whether these effects varied developmentally across childhood at the ages of 7, 9, and 12 years.

We found that the ADHD polygenic score (Demontis et al., 2019) was overall more predictive for behaviour problems than the neuroticism polygenic score (Luciano et al., 2018) except for with regard to emotional symptoms, explaining up to 2% of the variance. Similarly, environmental risk explained up to 2.7% of variance in behaviour problems and was on average more predictive for hyperactivity/inattention and conduct problems than for peer relationships and emotional symptoms. No clear developmental trends in these associations emerged across children's ages. Including GxE effects in the analyses improved the prediction of behaviour problems only slightly (Figure 4). Twelve out of 48 interaction tests were significant but they accounted on average only for 0.2% of variance, which was below the power threshold for detecting interaction effects in this sample. Despite their modest size, the interactions indicated that children with a higher genetic propensity for psychopathology showed disproportionally more problem behaviour when they experienced high levels of environmental risk. However, given the small effect sizes observed here, we caution against interpreting these findings with reference to any particular theoretical model of GxE. In order to be able to draw stronger conclusions from

these findings, GxE effects on children's behaviour problems need to be validated in future studies involving even larger sample sizes.




(c)
ADHD GPS
Neuroticism GPS
Environmental risk
Discipline
Age 7
Emotional symptoms
Age 9
Emotional symptoms
Age 12
Emotional symptoms
0.00
0.01
0.02
0.03
0.04
0.05

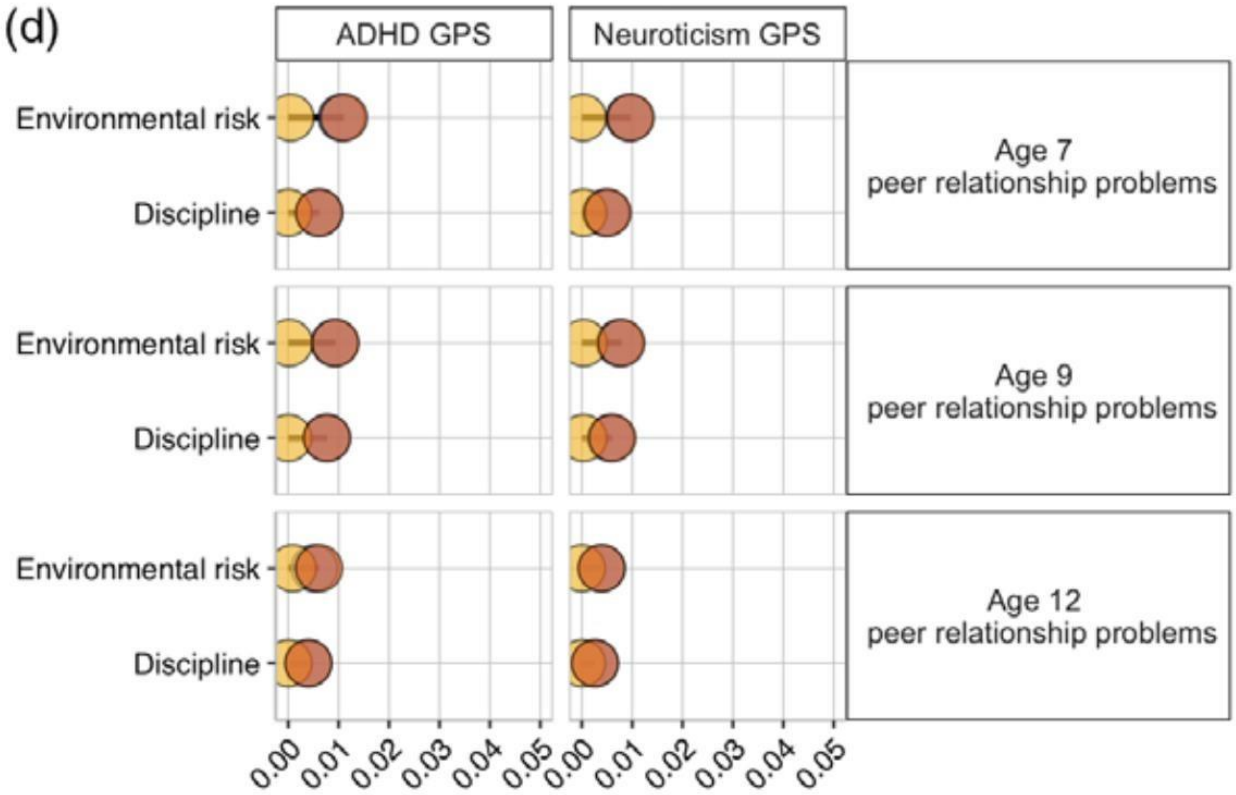

(d)
ADHD GPS
Neuroticism GPS
Environmental risk
Discipline
Age 7
peer relationship problems
Age 9
peer relationship problems
Age 12
peer relationship problems
0.00
0.01
0.02
0.03
0.04
0.05

Figure 4. G + E and GxE prediction of teacher-rated hyperactivity/inattention (a), conduct problems (b), emotional symptoms (c) and peer relationship problems (d) at ages 7, 9 and 12.

Note. Each line shows how much variance the environmental risk or discipline composite explains together with the polygenic score. Blue circles represent the additive prediction effects of environmental measures and the PGS, yellow circles show the interaction effects (controlling for the additive effects), and red shows the total variance explained by both additive and interaction effects. Adapted from Plomin et al. (2022).

### *Interim summary of empirical GxE findings*

We did not find conclusive evidence for gene-environment interaction effects in explaining children's cognitive and social-emotional development. If these findings reflected true (nonsignificant) effects, we could conclude that interventions which target children's cognitive and social-emotional development may be equally effective for all children that face challenges due to environmental risk and may therefore yield comparable results independent of child genotype. However, while this conclusion is supported by this project's empirical evidence, we caution that our null effects may be false negative findings for several reasons.

First, the polygenic scores currently available only explain a small proportion of variance in children's cognitive and social-emotional development and suffer from "missing heritability", meaning that these scores do not capture all trait-relevant DNA differences. The polygenic score for years spent in education (Lee et al., 2018) explains about 16% of the variation in educational achievement (GCSE results) and about 11% of variation in cognitive ability at age 16 (Allegrini et al., 2019; Wilding et al., 2023), while twin studies have confirmed heritability estimates of at least 50% (Polderman et al., 2015). One possible explanation is that polygenic scores only include more common DNA differences, and not those that are rare in the population. Some of these rarer variants might be just as predictive or perhaps even more predictive for trait outcomes than common DNA variants. Another possible explanation is that polygenic scores do not include gene-gene interactions such as dominance effects – that is, cases where the effect of one SNP on a trait outcome depends on another SNP; Manuck & McCaffery, 2014). Because interaction and main effects are independent of each other, the modest predictive validity of polygenic scores for children's differences in early life cognitive and social-emotional development does not necessarily cause null findings regarding potential interaction effects. However, as the polygenic score only captured a fraction of the DNA differences associated with

the trait of interest, true GxE effects may have been missed in the analyses we conducted here. It is also possible that our operationalisation of genetic propensities via polygenic scores might have not been suitable to identify GxE effects, because interactions might happen at a different level of measurement. That is, interaction effects may take place at the level of single genes or single SNPs. If this is the case, using polygenic scores, which are a composite measure, combining many SNPs together, would be too coarse and high-level to detect interactions.

Second, while we investigated a much broader range of environmental variables than previous studies in this field, there are many more that we did not test here. In the same way that traits are affected by many genetic variants across the genome, many environmental factors (e.g., family characteristics, parental behaviour, economic factors, educational provision, etc.) may influence observable trait outcomes. These measures broadly comprise the 'environome', a term coined to summarise all environmental factors that give rise to individual differences in cognitive and social-emotional development (von Stumm & d'Apice, 2022). However, there is no agreed-upon theory of how to map all possible environmental factors, rendering it impossible to know if all environmental domains were measured. It is possible, therefore, that GxE effects may exist in environments which were not measured, or that our results may have been affected by unobserved environments. Alternatively, our environmental measures might not have been well-suited in terms of the measurement level to detect GxE effects, in a similar way to our operationalisation of genetic propensities.

Taken together, the lack of GxE effects observed here may be due to methodological limitations of these papers, and GxE effects may be found if these limitations are overcome. Therefore, there is currently absence of evidence for GxE in children's cognitive and social-emotional development, and no true evidence for the absence of these interaction effects. Regarding our aforementioned research questions, we conclude that future studies on GxE in children's early development are needed to yield more conclusive results about the existence of interaction effects.

### *Why did our analyses not reveal robust GxE effects in children's development?*

The approach we took in the research outlined here does provide not produce strong, robust, conclusive evidence as to whether GxE effects are relevant for children's cognitive and social-emotional development. Results from previous research were also inconclusive (Hsu et al, 2012; Manuck & McCaffery, 2014). We therefore reviewed the methodological issues that might

have contributed to our null findings (von Stumm & Nancarrow, 2023). A crucial issue for identifying interaction effects is statistical power, which refers to the probability that a statistical test can detect a true effect (Abraham & Russell, 2008; Cohen, 1992). Underpowered studies lead to inconsistent results, as true effects are likely to go undetected, and effects which are detected may be artefacts (Abraham & Russell, 2008; Sham & Purcell, 2014). Because GxE effect sizes tend to be very small, detecting them in empirical studies requires large samples of thousands of participants, rendering many existing studies underpowered. No 'hard and fast' rule exists to calculate power for interaction studies (Shieh, 2009). We therefore performed simulation studies to estimate the required sample sizes needed to ensure sufficient statistical power (i.e., typically 80%) in order to detect GxE effects with different effect sizes ($\beta$ = 0.01 - 0.10). Our simulations suggested that a sample of over N = 75,000 is necessary to detect the smallest effects, while even for mid-sized interaction effects sample sizes of close to N = 4,000 are required (see Figure 5 for required sample sizes along the entire span of effect sizes modelled).

To inform future research on GxE effects, we have drawn several recommendations from our findings. For example, since the issue of statistical power is closely tied to a study's sample size, we recommend collecting data from larger numbers of participants or to increase sample sizes through collaborations and pooling of data. To realise such collaborations and the pooling of data, consistent data collection protocols need to be implemented, so that data can easily be harmonised and shared. Our recommendations on sample size, together with those on measurement quality and confounding, will provide a valuable resource for researchers studying GxE in child development, to detect true interaction effects and correctly interpret null results.

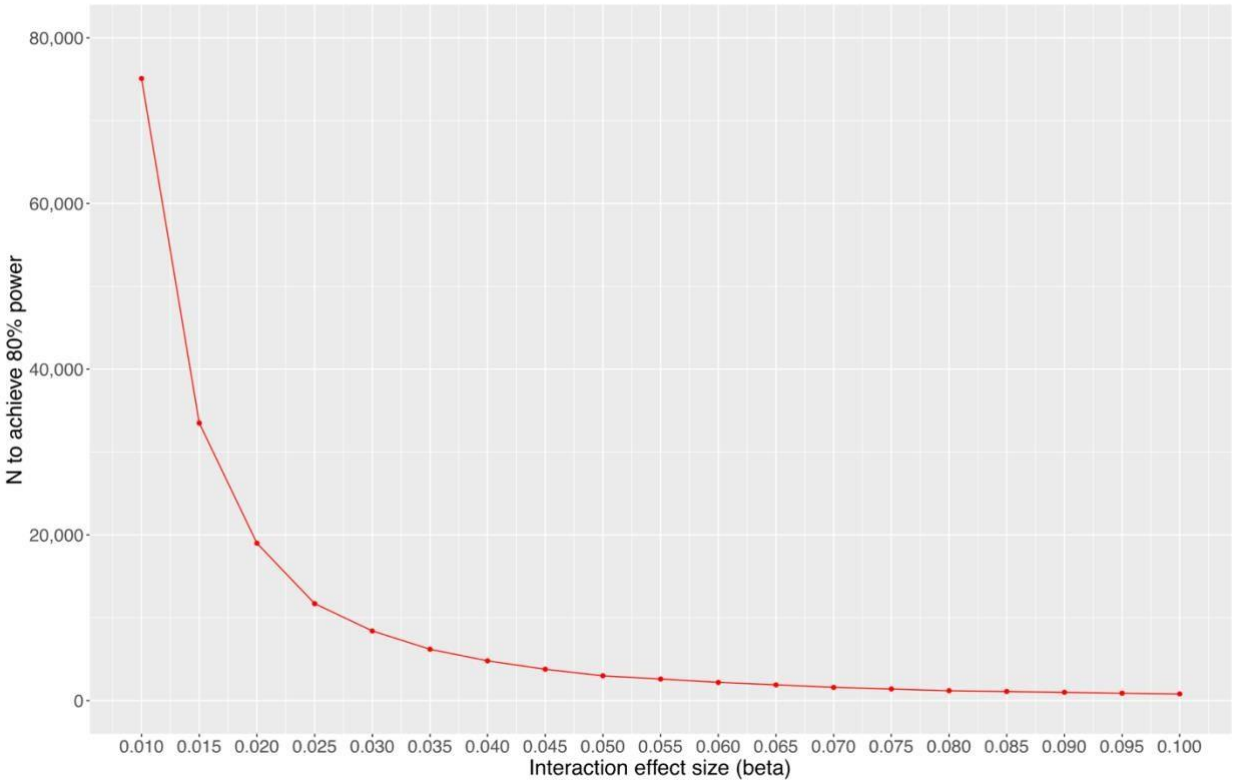



Figure 5. Sample sizes to achieve 80% statistical power for identifying interaction effect sizes from β = 0.01 to 0.10. Adapted from von Stumm and Nancarrow (2023).

## Discussion & Implications

### *Recap: Key findings and discussion points*

Our studies on GxE in children's early life development conducted over the course of this project did not identify robust GxE effects in the prediction of children's cognitive and social-emotional development. Both the wide variety of environmental measures included in our analyses and children's genetic propensities, operationalised as polygenic scores, explained significant portions of variance in children's developmental outcomes. However, over and above these direct, additive effects, interactions between genetic and environmental predictors do not add substantially to the variance explained.

Our null findings might indicate true absence of GxE in children's early life development. Alternatively, and more likely, our results may be false negatives that may have emerged for three main reasons. First, despite our preliminary power analyses, our studies might not have had sufficient statistical power to robustly identify the potentially miniscule and complex GxE effects in children's cognitive and social-emotional development. To explore this possibility, we conducted simulation studies on statistical power and reported the sample sizes needed to discover interactions of varying effect sizes. Second, the measures we used to operationalise environments and genetic propensities may have hindered detecting GxE effects, because they did not capture environments and genetic propensities at the right level to detect GxE. Since both, our genetic and environmental measures, can be seen as composites comprising information at a comparatively coarse level, we cannot rule out with certainty that GxE effects may exist at a more fine-grained level, that is, at the level of single genes or DNA variants and specific, narrow environmental factors. Third, our null findings may have resulted from the fact that the mechanisms driving gene-environment interplay are highly complex. We therefore investigated the bidirectional developmental relationship between children's cognitive development and their environmental experiences, in particular the cognitive stimulation they receive in their family homes in early life. We found bidirectional effects between cognitive stimulation and cognitive development, which were explained by common genetic and shared environmental factors. This finding illustrates that children are not passive recipients of the environments that they are exposed to. Instead, they actively select, shape, and create the environmental experiences that match their genetic propensities.

*Implications and impact*

While no robust evidence for GxE in children's early development emerged from our studies, we have confirmed the prediction from a wide variety of environmental factors and children's genetic propensities for cognitive and social-emotional development in the first years of life. From these findings, we can draw several implications for theory and practice, which are key to advancing the development of interventions that effectively ameliorate the pervasive effects of childhood inequality.

In general, the ability to systematically predict children's developmental differences, indicates the possibility to reliably identify children who are at an increased risk for poor developmental outcomes. Reliable identification of at-risk individuals is key to intervening and reducing the consequences of early life disadvantage (von Stumm, 2022). Our studies have revealed two principles that drive the relationship between environmental factors and children's developmental outcomes. First, we have found that many environmental factors, each with very small effects, seem to be involved in shaping child development, instead of just a handful of factors exerting stronger influences. This finding reflects the genetic principle of polygenicity – the phenomenon whereby each trait is influenced by a large number of genetic variants with miniscule effects that add up and, in sum, predict significant proportions of variance in traits. The same seems to be true for environmental factors when considering that up to 39 environmental predictors together account for about 20% of the variance in children's cognitive development (von Stumm et al., 2023). Second, these environmental factors most likely do not predict outcomes in one specific developmental domain, but are likely associated with children's development across domains, i.e., across both cognitive and social-emotional outcomes. This is akin to the principle of *pleiotropy* in genetics – the phenomenon whereby one genetic variant is associated with differences in trait expression for more than one phenotype.

Regarding the environment, these principles imply that changing a single or very few environmental factors may only have a rather small effect on development. Yet, at the same time, these changes to a single or very few environmental factors are likely to affect outcomes in more than one developmental domain, if they have an effect at all, even though their effect sizes are likely to be small. For example, regarding the bidirectional relationship we observed between children's cognitive development and the cognitive stimulation they receive, we speculate that increasing cognitive stimulation will be associated, not only with children's cognitive abilities, but also their social-emotional developmental outcomes, such as sociability and self-control. However, the link between cognitive stimulation and development should not be

deemed causal, because it rather is accounted for by common genetic and shared environmental factors. Further research should explore other aspects in the wide range of other putatively environmental factors, which can be expected to also play a role.

Our findings contribute to the development of realistic expectations for the effectiveness of interventions that seek to reduce early childhood inequality. Interventions are unlikely to produce large effects that are limited to one or two developmental domains; instead, they are likely to create very small changes which are spread broadly across developmental domains, if they are indeed successful. Recognising this is key to realistically predicting what an intervention can and cannot achieve. While interventions are typically costly to design and implement, upholding these efforts is crucial to eventually reduce the long-term effects of experiencing early life disadvantages. Recognising the potential for positive change and the limitations of interventions will help to maintain public support for continuing to develop and implement interventions, even if their benefits unfold only slowly and step by step (von Stumm, 2022).

## Recommendations for future research

### *Overcoming missing heritability: capturing more genetic differences*

As discussed above, our null GxE effects may be false negative results due methodological limitations of our work. One of these limitations is the problem of missing heritability. To ensure that possible GxE effects do not go unobserved, future research should strive to close the missing heritability gap. A possible solution to this problem could be using polygenic scores from larger GWAS that are sufficiently powered to detect more of the SNPs that are associated with the trait of interest, including SNPs that occur less frequently across individuals in a given population. A GWAS for human height with a sample size of N > 5,000,000 recently became the first GWAS to overcome the missing heritability gap, proving that this goal is attainable (Yengo et al, 2022). Larger GWASs are also presently being conducted for psychologically relevant traits, such as educational achievement, and efforts to pool samples and harmonize measures are increasing. This means that future studies will be able to utilise polygenic scores with higher predictive validity. Additionally, future studies could refer to methods that combine GWAS for different traits (such as genomic structural equation modelling; Grotzinger et al, 2019) to improve the predictive ability of polygenic scores.

If, alternatively, the reason was for our null findings was that polygenic scores represent too coarse a level of measurement, then using stronger polygenic scores in future analyses may not overcome this limitation. However, in this case, a larger GWAS, identifying more of the SNPs associated with a trait of interest would also enable researchers to search for a larger set of gene-environment interactions at the more fine-grained SNP-level.

### *Studying the Environome: capturing more environmental differences*

Complex traits such as child development are associated with many environmental factors across the 'environome'. However, there are no guidelines or methods for assessing the environome, akin to those for assessing a person's genome (von Stumm & d'Apice, 2022). Therefore, GxE effects may be missed if environmental domains are not studied comprehensively, but instead through cumulative risk indices. Future research should therefore be targeted at collecting naturalistic observations in real time, for instance by using digital technologies such as mobile sensing for "capturing the environome across levels, dimensions, and time in unprecedented depth and detail" (von Stumm & d'Apice, 2022, p. 6). This is needed to develop a scientific framework of environmental domains, and to investigate GxE in children's development across these domains. By observing finer-grained individual differences in genetic and environmental factors, future research can determine where GxE effects exist in early life development.

### *The utility of polygenic scores in intervention research*

The key motivation for studying GxE in children's cognitive and social-emotional development is to inform interventions that improve children's developmental trajectories. The effectiveness of a given intervention may depend on participating children's genotype if the environment that is targeted by the intervention is subject to GxE effects. The effectiveness of interventions varies substantially across children (Jeong et al, 2021; Protzko et al, 2013). To better understand how genetics may contribute to children's differences in responding to interventions, study designs for evaluating the effectiveness of interventions should be genetically sensitive. That is, such studies should include genetic propensities as predictors or confounders alongside their other variables. Polygenic scores offer an unprecedented opportunity for integrating genetic effects in intervention studies, because they are easily obtained and can be modelled like any other predictor variable. GWAS summary statistics are openly available, and genotyping samples has

become increasingly affordable, facilitating the inclusion of genetics, in the form of polygenic scores, into studies that previously relied solely on phenotypic and environmental measures of children's developmental differences. The methodological accomplishment of polygenic scores – perhaps the epitome of the DNA revolution – will not only advance our understanding of children's genetic propensities, but strengthen discoveries of environmental influences as previously unobserved, confounding genetic factors can be accounted for. We predict that the advent of more robust findings on the gene-environment interplay in children's early life development is imminent, and that intervention researchers who incorporate polygenic scores in their study designs will be key drivers of this process.

## Conclusions

The aim of our project was to investigate GxE effects in children's early life cognitive and social-emotional development, since previous findings in this area have been inconclusive. While no robust findings on GxE emerged from our analyses, our findings contribute to improving our understanding of how genetic and environmental factors jointly predict differences in children's cognitive and social-emotional development. We identified a wide range of genetic and environmental factors, that – each with small effects – are likely to drive children's differences over the course of development. The influences of these genetic and environmental factors are neither specific to a particular phenotypic trait or developmental outcome, nor are they independent of each other. Overlapping or correlated sets of genetic and environmental factors drive children's developmental differences across domains, with small individual effect sizes. We recommend that future research on GxE places particular emphasis on statistical power and seeks large sample sizes that allow for reliably detecting small interaction effects.

## References


Abraham, W. T. and Russell, D. W. (2008), Statistical power analysis in psychological research. *Social and Personality Psychology Compass, 2,* 283-301. https://doi.org/10.1111/j.1751-9004.2007.00052.x

Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry*, *24*(6), Article 6. https://doi.org/10.1038/s41380-019-0394-4

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.). Psychological Corporation.

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

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., ... & Lancet Early Childhood Development Series Steering Committee. (2017). Early childhood development coming of age: science through the life course. *The Lancet, 389(10064),* 77-90. https://doi.org/10.1016/S0140-6736(16)31389-7

Blaggan, S., Guy, A., Boyle, E. M., Spata, E., Manktelow, B. N., Wolke, D., & Johnson, S. (2014). A Parent Questionnaire for Developmental Screening in Infants Born Late and Moderately Preterm. *Pediatrics*, *134*(1), e55–e62. https://doi.org/10.1542/peds.20140266

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry*, *176*(5), 376-387. https://doi.org/10.1176/appi.ajp.2018.18070881

Bricker, D., Squires, J., & Mounts, L. (1995). *Ages and stages questionnaire: A parentcompleted, child-monitoring system*. Brookes.

Bronfenbrenner, U., & Morris, P. A. (2006). The Bioecological Model of Human Development. In *Handbook of child psychology: Theoretical models of human development, Vol. 1, 6th ed* (pp. 793–828). John Wiley & Sons Inc.

Cattan, S., Fitzsimons, E., Goodman, A., Phimister, A., Ploubidis, G. B., and Wertz, J. (2022), 'Early Childhood Inequalities', IFS Deaton Review of Inequalities.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M., Johannesson, M., Magnusson, P. K. E., Lichtenstein, P., Atwood, C. S., Freese, J., Hauser, T. S., Hauser, R. M., Christakis, N., & Laibson, D. (2012). Most Reported Genetic Associations With General Intelligence Are Probably False Positives. *Psychological Science*, *23*(11), 1314–1323. https://doi.org/10.1177/0956797611435528

Cheesman, R., Borgen, N.T., Lyngstad, T.H. *et al.* A population-wide gene-environment interaction study on how genes, schools, and residential areas shape achievement. *npj Sci. Learn.* **7**, 29 (2022). https://doi.org/10.1038/s41539-022-00145-8

Chen, S., & Parmigiani, G. (2007). Meta-analysis of BRCA1 and BRCA2 penetrance. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, *25*(11), 1329. https://doi.org/10.1200/JCO.2006.09.1066

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

Cohen, J. (1992a). Statistical Power Analysis. *Current Directions in Psychological Science*, *1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Cohen, J. (1992b). A power primer. Psychological Bulletin, 112(1), 155-159. https://doi.org/10.1037//0033-2909.112.1.155

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry*, *150*(6), 782–786. https://doi.org/10.1192/bjp.150.6.782

Culverhouse, R. C., Saccone, N. L., Horton, A. C., Ma, Y., Anstey, K. J., Banaschewski, T., ... & Bierut, L. J. (2018). Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-HTTLPR genotype contributing to the development of depression. *Molecular psychiatry*, *23*(1), 133-142. https://doi.org/10.1038/mp.2017.44

d'Apice, K., Latham, R. M., & von Stumm, S. (2019). A naturalistic home observational approach to children's language, cognition, and behavior. *Developmental Psychology*, *55*, 1414–1427. https://doi.org/10.1037/dev0000733

Deary IJ. Intelligence. Annu Rev Psychol. 2012;63:453-82. doi: 10.1146/annurev-psych-120710100353. Epub 2011 Sep 19. PMID: 21943169.

Deater-Deckard, K. (2000). Parenting and child behavioral adjustment in early childhood: A quantitative genetic approach to studying family processes. *Child Development*, *71*(2), 468–484. https://doi.org/10.1111/1467-8624.00158

Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology*, *10*(3), 469–493. https://doi.org/10.1017/s0954579498001709

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

DfEE, (Department for Education & Employment). (2000). *[ARCHIVED CONTENT]*. https://webarchive.nationalarchives.gov.uk/ukgwa/20130320225719/https://media.education.gov.uk/assets/files/pdf/sfr092000pdf.pdf

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology, 7, 383-409.* https://doi.org/10.1146/annrev-clinpsy032210-104518

Duckworth, A., & Gross, J. J. (2014). Self-Control and Grit: Related but Separable Determinants of Success. *Current Directions in Psychological Science*, *23*(5), 319–325. https://doi.org/10.1177/0963721414541462

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual* (2nd ed.). Brookes.

Gidziela, A., Rimfeld, K., Malanchini, M., Allegrini, A. G., McMillan, A., Selzam, S., Ronald, A., Viding, E., von Stumm, S., Eley, T. C., & Plomin, R. (2022). Using DNA to predict behaviour problems from preschool to adulthood. *Journal of Child Psychology and Psychiatry*, *63*(7), 781–792. https://doi.org/10.1111/jcpp.13519

Goodman, A., Joshi, H., Nasim, B., & Tyler, C. (2015*). Social and emotional skills in childhood and their long-term effects on adult life.* London: Institute of Education. https://www.eif.org.uk/report/social-and-emotional-skills-in-childhood-and-their-longterm-effects-on-adult-life

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586. https://doi.org/10.1111/j.14697610.1997.tb01545.x

Greven, C. U., Lionetti, F., Booth, C., Aron, E. N., Fox, E., Schendan, H. E., ... & Homberg, J. (2019). Sensory processing sensitivity in the context of environmental sensitivity: A critical review and development of research agenda. *Neuroscience & Biobehavioral Reviews, 98,* 287-305. https://doi.org/10.1016/j.neubiorev.2019.01.009

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525. https://doi.org/10.1038/s41562-019-0566-x

Harden, K. P. (2021). “Reports of my death were greatly exaggerated”: Behavior genetics in the postgenomic era. *Annual Review of Psychology, 72,* 37-60. https://doi.org/10.1146/annurev-psych-052220-103822

Hettema, J. M., Neale, M. C., & Kendler, K. S. (2001). A review and meta-analysis of the genetic epidemiology of anxiety disorders. *American Journal of Psychiatry, 158(10),* 1568-1578 https://doi.org/10.1176/appi.ajp.158.10.1568

Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., ... & McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience, 22(3),* 343-352. https://doi.org/10.1038/s41593-018-0326-7

Hsu, L., Jiao, S., Dai, J. Y., Hutter, C., Peters, U., & Kooperberg, C. (2012). Powerful Cocktail Methods for Detecting Genome-Wide Gene-Environment Interaction. *Genetic Epidemiology*, *36*(3), 183–194. https://doi.org/10.1002/gepi.21610

Ireton, H., & Thwing, E. (1974). *The Minnesota child development inventory*. Behavior Science Systems.

Jeong, J., Franchett, E. E., Ramos de Oliveira, C. V., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS medicine, 18*(5), e1003602. https://doi.org/10.1371/journal.pmed.1003602

Kwon, J. M., & Goate, A. M. (2000). The candidate gene approach. *Alcohol research & health: the journal of the National Institute on Alcohol Abuse and Alcoholism*, *24*(3), 164–168. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6709736/

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

Luciano, M., Hagenaars, S. P., Davies, G., Hill, W. D., Clarke, T.-K., Shirali, M., Harris, S. E., Marioni, R. E., Liewald, D. C., Fawns-Ritchie, C., Adams, M. J., Howard, D. M., Lewis, C. M., Gale, C. R., McIntosh, A. M., & Deary, I. J. (2018). Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nature Genetics*, *50*(1), Article 1. https://doi.org/10.1038/s41588-017-0013-8

Manuck, S. B., & McCaffery, J. M. (2014). Gene-environment interaction. *Annual Review of Psychology*, *65*, 41–70. https://doi.org/10.1146/annurev-psych-010213-115100

Martin, A. J., Darlow, B. A., Salt, A., Hague, W., Sebastian, L., McNeill, N., & Tarnow-Mordi, W. (2013). Performance of the Parent Report of Children's Abilities–Revised (PARCA-R) versus the Bayley Scales of Infant Development III. *Archives of Disease in Childhood*, *98*(12), 955–958. https://doi.org/10.1136/archdischild-2012-303288

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of*

*Applied Developmental Psychology*, *16*(3), 429–444. https://doi.org/10.1016/01933973(95)90028-4

McCarthy, D. (1972). *McCarthy scales of Children's abilities.* The Psychological Corporation.

Moore, D. S., & Shenk, D. (2017). The heritability fallacy. Wiley Interdisciplinary Reviews: *Cognitive Science, 8(1-2),* e1400. https://doi.org/10.1002/wcs.1400

NICE, N. I. for C. E. (2017). *Developmental follow-up of children and young people born preterm, NICE guideline [NG72]*. NICE. https://www.nice.org.uk/guidance/ng72/chapter/Recommendations#providing-enhanceddevelopmental-support

Oginni, O., Starr, A., & von Stumm, S. (2024). *Do children cause the cognitive stimulation they receive? Modelling the direction of causality*. PsyArXiv. https://doi.org/10.31234/osf.io/pqf78

Oliver, B., Dale, P., Saudino, K., Petrill, S., Pike, A., & Plomin, R. (2002). The Validity of a Parent-based Assessment of Cognitive Abilities in Three-year Olds*. *Early Child Development and Care*, *172*(4), 337–348. https://doi.org/10.1080/03004430212713

ONS, (Office for National Statistics). (2000). *SOC 2000—Previous version of the Standard Occupational Classification*. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/socarchive

ONS, (Office for National Statistics). (2017). *More mothers with young children working fulltime—Office for National Statistics*. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/moremotherswithyoungchildrenworkingfulltime/2017-09-26

Oxley, F. A. R., Wilding, K., & von Stumm, S. (2023). DNA & IQ: Big deal or much ado about nothing?. https://osf.io/63zmr/

Pingault, J., Allegrini, A. G., Odigie, T., Frach, L., Baldwin, J. R., Rijsdijk, F., & Dudbridge, F. (2022). Research Review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *Journal of Child Psychology and Psychiatry*, *63*(10), 1125–1139. https://doi.org/10.1111/jcpp.13607

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

Plomin, R., & von Stumm, S. (2021). Polygenic scores: prediction versus explanation. *Molecular Psychiatry* 27, 49–52. https://doi.org/10.1038/s41380-021-01348-y

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

Plomin, R., Gidziela, A., Malanchini, M., & Stumm, S. von. (2022). Gene–environment interaction using polygenic scores: Do polygenic scores for psychopathology moderate predictions from environmental risk to behavior problems? *Development and Psychopathology*, *34*(5), 1816–1826. https://doi.org/10.1017/S0954579422000931

Pluess, M., & Belsky, J. (2013). Vantage sensitivity: Individual differences in response to positive experiences. *Psychological Bulletin, 139*(4), 901–916. https://doi.org/10.1037/a0030196

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

Protzko, J., Aronson, J., & Blair, C. (2013). How to Make a Young Child Smarter: Evidence From the Database of Raising Intelligence. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *8*(1), 25–40. https://doi.org/10.1177/1745691612462585

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*(1), 149-150. https://doi.org 10.1093/bioinformatics/19.1.149

Rimfeld, K., Malanchini, M., Spargo, T., Spickernell, G., Selzam, S., McMillan, A., Dale, P. S., Eley, T. C., & Plomin, R. (2019). Twins Early Development Study: A genetically sensitive investigation into behavioural and cognitive development from infancy to emerging adulthood. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, *22*(6), 508–513. https://doi.org/10.1017/thg.2019.56

Saudino, K. J., Dale, P. S., Oliver, B., Petrill, S. A., Richardson, V., Rutter, M., Simonoff, E., Stevenson, J., & Plomin, R. (1998). The validity of parent-based assessment of the cognitive abilities of 2-year-olds. *British Journal of Developmental Psychology*, *16*(3), 349–362. https://doi.org/10.1111/j.2044-835X.1998.tb00757.x

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., … Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, *50*(7), Article 7. https://doi.org/10.1038/s41588-018-0152-6

Schmitz, L. L., & Conley, D. (2017). The effect of Vietnam-era conscription and genetic potential for educational attainment on schooling outcomes. *Economics of Education Review*, *61*, 85–97. https://doi.org/10.1016/j.econedurev.2017.10.001

Sham, P. C., & Purcell, S. M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics*, *15*(5), Article 5. https://doi.org/10.1038/nrg3706

Shieh, G. (2009). Detecting Interaction Effects in Moderated Multiple Regression With Continuous Variables Power and Sample Size Considerations. *Organizational Research Methods*, *12*(3), 510–528. https://doi.org/10.1177/1094428108320370

Thomas, D. (2010). Gene–environment-wide association studies: emerging approaches. Nature Reviews Genetics 11, 259–272. https://doi.org/10.1038/nrg2764

Tucker-Drob, E. M., & Bates, T. C. (2016). Large Cross-National Differences in Gene × Socioeconomic Status Interaction on Intelligence. *Psychological Science*, *27*(2), 138–149. https://doi.org/10.1177/0956797615612727

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science*, *9*(5), 160–164. https://doi.org/10.1111/14678721.00084

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), Article 1. https://doi.org/10.1038/s43586-021-00056-9

van de Weijer, M. P., Pelt, D. H. M., de Vries, L. P., Baselmans, B. M. L., & Bartels, M. (2022). A Re-evaluation of Candidate Gene Studies for Well-Being in Light of Genome-Wide Evidence. *Journal of Happiness Studies*, *23*(6), 3031–3053. https://doi.org/10.1007/s10902-022-00538-x

von Stumm, S., Smith-Woolley, E., Ayorech, Z., McMillan, A., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Predicting educational achievement from genomic measures and socioeconomic status. *Developmental Science, 23*(3). https://doi.org/10/1111/desc.19925

von Stumm, S. (2022). Early childhood inequalities: The rocky path from observation to action. *IFS Deaton Review of Inequalities*. https://ifs.org.uk/inequality/early-childhoodinequalities-the-rocky-path-from-observation-to-action

von Stumm, S., & d'Apice, K. (2022). From Genome-Wide to Environment-Wide: Capturing the Environome. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *17*(1), 30–40. https://doi.org/10.1177/1745691620979803

von Stumm, S., Kandaswamy, R., & Maxwell, J. (2023). Gene-environment interplay in early life cognitive development. *Intelligence*, *98*, 101748. https://doi.org/10.1016/j.intell.2023.101748

von Stumm, S., & Nancarrow, A. F. (2023). *New methods, persistent issues, and one solution: Gene-environment interaction studies of childhood cognitive development* [Preprint]. https://doi.org/10.31235/osf.io/y28h6

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

Wilding, K., Wright, M., & von Stumm, S. (2023). *Using DNA to predict Educational Success: A Meta-Analytic review*. PsyArXiv. https://doi.org/10.31234/osf.io/f7vt9

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics, 88*(1), 76-82

Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliasen, A. U., Jiang, Y., Raghavan, S., Miao, J., Arias, J. D., Graham, S. E., Mukamel, R. E.,

Spracklen, C. N., Yin, X., Chen, S.-H., Ferreira, T., Highland, H. H., … Hirschhorn, J. N. (2022). A saturated map of common genetic variants associated with human height. *Nature*, *610*(7933), 704–712. https://doi.org/10.1038/s41586-022-05275-y

Zuckerman, M. (1999). Diathesis-stress models. In M. Zuckerman, *Vulnerability to psychopathology: A biosocial model* (pp. 3–23). American Psychological Association. https://doi.org/10.1037/10316-001