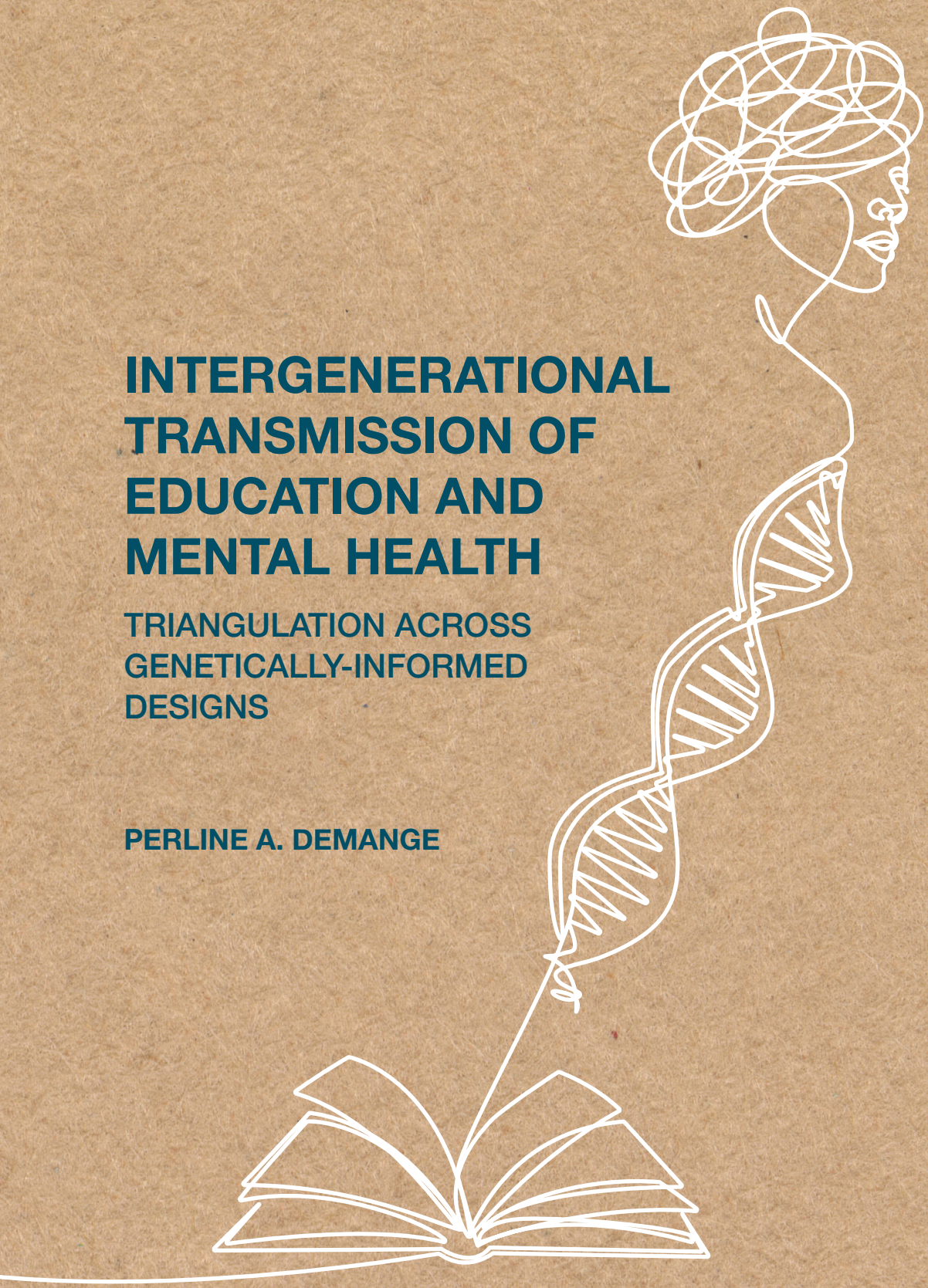
INTERGENERATIONAL TRANSMISSION OF EDUCATION AND MENTAL HEALTH

TRIANGULATION ACROSS GENETICALLY-INFORMED DESIGNS

PERLINE A. DEMANGE



INTERGENERATIONAL TRANSMISSION OF EDUCATION AND MENTAL HEALTH

TRIANGULATION ACROSS GENETICALLY-INFORMED DESIGNS

PERLINE A. DEMANGE

Paranymphs: Anne M. Hendriks
Nicole Huizinga

Cover and layout: Céline Marchal
Print: Ridder print | www.ridderprint.nl
ISBN: 978-94-6483-358-4
DOI: 10.5463/thesis.277

VRIJE UNIVERSITEIT

INTERGENERATIONAL TRANSMISSION OF EDUCATION AND MENTAL HEALTH: TRIANGULATION ACROSS GENETICALLY-INFORMED DESIGNS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op maandag 2 oktober 2023 om 9.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Perline Aline Delle Demange

geboren te Nancy, Frankrijk

promotor:	prof.dr. D.I. Boomsma
copromotoren:	dr. M.G. Nivard dr. E. van Bergen
promotiecommissie:	prof.dr. D. Posthuma prof.dr. H.G. van de Werfhorst prof.dr. P.A. Dykstra prof.dr. J.L.W. van Kippersluis dr. L.W. Wesseldijk dr. N. van den Berg

TABLE OF CONTENTS

Chapter 1. General introduction 9

Chapter 2. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction 19

Published as: **Demange, P.A.***, Malanchini, M.*, Mallard, T.T., Biroli, P., Cox, S.R., Grotzinger, A.D., Tucker-Drob, E.M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D.I., Caspi, A., Corcoran, D.L., Domingue, B. W., Harris, K.M., Ip, H.F., Mitchell, C., Moffitt, T.E., Poulton, R., Prinz, J.A., Sugden, K., Wertz, J., Williams, B.S., de Zeeuw, E.L., Belsky, D.W.*, Harden, K.P.* & Nivard, M.G.* (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. Nature Genetics, 53(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Chapter 3. Estimating effects of parents' cognitive and noncognitive skills on offspring education using polygenic scores 45

Published as: **Demange, P.A.**, Hottenga, J.J., Abdellaoui, A., Eilertsen, E.M., Malanchini, M., Domingue, B.W., Armstrong-Carter, E., de Zeeuw, E.L., Rimfeld, K., Boomsma D.I., van Bergen, E., Breen, G., Nivard, M.G., & Cheesman, R. (2022). Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. Nature Communications 13, 4801 https://doi.org/10.1038/s41467-022-32003-x

Chapter 4. Evaluating the causal relationship between educational attainment and mental health 81

Under review and preprinted as: **Demange, P.A.**, Boomsma, D.I., van Bergen, E.*, Nivard, M.G.* (2023) Evaluating the causal relationship between educational attainment and mental health. medRxiv, 2023.01.26. https://doi.org/10.1101/2023.01.26.23285029

Chapter 5. No effect of parental mental health on children's academic achievement: a within-family study 105

In preparation as: **Demange, P.A.**, Nivard, M.G., Torvik, F., Eilertsen, E.M., Cheesman. R., Nordmo, M., Lyngstad, T., Ystrøm, E.*, van Bergen, E.* No effect of parental mental health on children's academic achievement: a within-family study

Chapter 6. Summary of my research chapters 119

Chapter 7. General discussion and future directions 127

References 142
Author contributions & competing interests 166
List of publications 168
Acknowledgements 172

01

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1. GENERAL INTRODUCTION

There is a well-established association between the level of education attained by parents and their children[1]. This intergenerational transmission of educational attainment is a phenomenon observed since the 19th century, prevalent in most countries, and strongly linked to the transmission of broader socioeconomic status (SES)[1–5]. Many other traits have been found to be correlated across generations, including mental health indicators. Offspring of parents with mental health issues are at an increased risk of developing such issues themselves[6]. Furthermore, there is an unequal educational distribution of mental illness, with individuals with lower levels of education being more likely to exhibit symptoms and to be diagnosed, and individuals with mental illness being more likely to attain lower levels of education[7]. However, the underlying causal mechanisms of these relationships remain poorly understood.

In this doctoral thesis, I aim to deepen our understanding of the mechanisms underlying the correlations between educational attainment and mental health, and the intergenerational transmission of these traits. The set of (intergenerational) relationships that might exist can be found in Figure 1. To investigate the highlighted relationships, I employ genetics as a tool. Genetically-informed designs exploit genetic information inferred from measured genetic variation (difference in DNA between people) and/or from the genetic similarity expected from familial relationships. These designs are uniquely qualified to help to unveil these mechanisms[8–12]: they offer new approaches to investigate the aetiology of traits and they represent powerful tools to account for genetic and environmental confounding, enhancing causal inference.

When possible, I triangulate across datasets and across various complementary research designs. Triangulation is the practice of integrating results from multiple sources of data, methodological approaches, or multiple analytical designs, thus reducing the impact of biases, errors, or limitations that may arise from relying on a single method or source of data. By examining the same phenomenon using several approaches, triangulation increases the reliability of our findings and provides a more comprehensive understanding of the research question at hand[13–16].

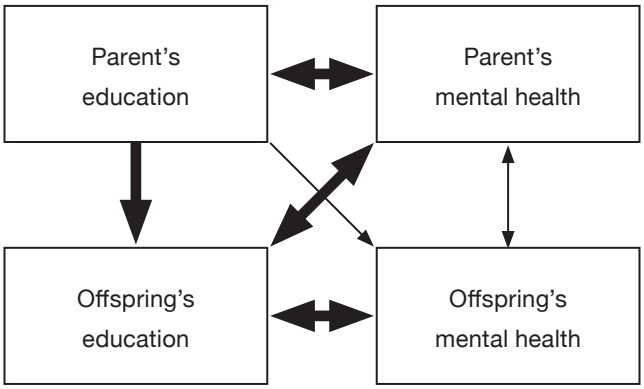


Figure 1 | Hypothesized intergenerational effects of education and mental health. Double-headed arrows are pathways for which I do not have a strong hypothesis on the direction of causation, or for which bidirectional effects are hypothesized. Single-headed arrows are pathways for which a direction of causation is hypothesized: parents' education is only believed to be affected by their offspring's education or mental health in the rare case of parents being in education during or after the child is of school age. In this thesis, I investigate the highlighted pathways.

By using triangulation across genetically-informed designs, I hope to strengthen causal inferences and provide a more comprehensive understanding of educational success, and its link with mental health, as well as the factors involved in the transmission of these traits across generations. I investigate what contributes to educational attainment, notably cognitive and noncognitive aspects of educational attainment **(Chapter 2)**, and how genetically-informed designs might reveal parental environmental effects on their offspring's education **(Chapter 3)**. I investigate the causal relationship between educational attainment and mental health **(Chapter 4)**, and the effect of parental mental health on their children's academic achievement **(Chapter 5)**.

Noncognitive aspects of educational attainment

Many individual traits might contribute to the education success of one individual, especially regarding their educational attainment. Educational *attainment* is the level of education a person has completed, while educational *achievement* refers to academic performance, used as a measure of the knowledge and abilities acquired during education. One well-studied trait contributing to education is cognitive performance (i.e. cognitive skills or intelligence), as measured by cognitive tests. Cognitive performance is known to be correlated to educational achievement and attainment between 0.4 and 0.8[17–20], sharing large overlapping genetic influences (genetic correlations ~ 0.7)[21–23]. Though it should be noted the relation between cognition and education is likely bidirectional[24], higher cognitive performance increases educational opportunities. However, traits beyond cognitive performance also contribute

to educational attainment. These traits have been named according to what they are not: noncognitive skills. They are also sometimes referred to as soft skills or socio-emotional competencies. **But what are these noncognitive skills? Do they matter for educational attainment and other life outcomes?**

Noncognitive skills suspected to matter for education include self-control, persistence, attention, motivation, curiosity, and consciousness[25–29]. So far, research has not reached a consensus on a definition of these noncognitive aspects of education. These factors have been measured and defined in many ways across different individual's ages, cohorts and scientific fields. Different measures differ in their consistency over time, in how heritable they are and, notably, they differ in their association with educational outcomes[25,30].

Genetic analyses can help us clarify the contribution of noncognitive skills to educational attainment and what these skills are and/or how they relate to other traits. A genome-wide association study (GWAS) examines the association of millions of genetic variants (single nucleotide polymorphisms: SNPs) with a trait. This information allows an understanding of the genetic architecture of the trait, including which genes and biological pathways might be involved in the expression of this trait. GWAS data makes it possible to calculate genetic correlations between any two GWAS-ed traits, even when they have not been measured in the same sample. A GWAS of noncognitive aspects of educational attainment could therefore help us clarify the genetic aetiology of noncognitive skills and their contribution to the (genetic) correlations between education attainment and other behavioural and health outcomes.

There has been an increase in our understanding of the genetic aetiology of cognitive skills and education attainment in the last years, whereas the genetics of noncognitive aspects of education are still mostly unknown. As GWAS requires the measure of the target trait in a large genotyped sample, the heterogeneity and lack of consistent noncognitive measures across large cohorts with genetic data prohibit a direct GWAS of noncognitive skills.

In **Chapter 2,** I develop a new approach to run a GWAS, which we named "GWAS-by-subtraction". I apply this approach to identify genetic variants associated with noncognitive aspects of educational attainment. As it is difficult to obtain noncognitive measures in large samples, I rely on previously published GWAS data and develop a method that allows to study a phenotype that was never directly measured. As the nature of noncognitive skills is unclear, noncognitive skills were conceptualized as the broad phenotype it was first defined as: all contributions to educational attainment that are not cognitive skills[31,32]. This GWAS-by-subtraction relies on structural equation modelling to subtract the genetic variance of cognitive performance from the genetic variance of educational attainment, giving us access

to the genetic variance of noncognitive aspects of educational attainment. This GWAS of noncognitive aspects of educational attainment should therefore be highly associated with educational attainment but not with cognitive performance.

To perform the GWAS-by-subtraction, well-powered GWASs of educational attainment (EA; N = 1.1M)[21] and of cognitive performance (N = 257,841)[21], executed in European-ancestry populations, were available. To confirm the GWAS of noncognitive aspects of EA is highly associated with EA but not with cognitive performance, I investigate its associations in out-of-sample data, including the Netherlands Twin Registry, the National Longitudinal Study of Adolescent to Adult Health, the Texas Twins Project, the Dunedin Longitudinal Study, E-Risk Longitudinal Twin Study, and the Wisconsin Longitudinal Study cohorts. Using cognitive and educational measures in these cohorts and existing GWAS summary statistics, I perform a phenotypic annotation to identify behavioural, psychological, and health traits associated with the cognitive and/or noncognitive components of education. Using several gene-expression datasets and brain imaging GWASs, biological annotation was done to identify shared and specific neurobiological correlates of the cognitive and noncognitive components of education.

Parental noncognitive influences on offspring's education

Parents and their offspring tend to have similar educational outcomes. Many studies have investigated how much certain parental characteristics might influence offspring education, but relatively few have considered noncognitive aspects of educational attainment. **Do parents' noncognitive skills affect their offspring's educational outcomes?** Research suggests that parents socially influence their children's noncognitive skills including emotion regulation, social capacities, attitudes and motivations[33,34]. Given that noncognitive skills seems to support education, we expect parents' noncognitive skills to also affect children's educational outcomes possibly through parenting and creating a nurturing environment.

Answering this question is challenging. As discussed for **Chapter 2**, the assessment of parental noncognitive skills is not straightforward. An additional challenge is the need to distinguish the social transmission of skills from parent to child from genetic transmission. Parents not only shape their children's environment, but they also pass on their genetic predispositions, which makes it difficult to distinguish the effects of one from the other[35]. We cannot simply correlate parental noncognitive skills to child school outcomes to establish a causal relation.

Genetically-informed designs can help us assess the potential environmental effect of parental noncognitive skills, controlling for genetic transmission. Relying on the GWAS

of noncognitive aspects of educational attainment developed in **Chapter 2**, I compute polygenic scores (PGS) of cognitive and of noncognitive aspects of EA. By aggregating the effects of one person's genetic variants based on the effect sizes from the GWAS, I obtain an index summarizing the individual's trait-specific genetic endowment. I therefore obtain a proxy measure of (the genetics of) noncognitive skills in large cohorts with genotyped individuals. While this proxy measure is imperfect, it is comparable across cohorts and allows to obtain information for individuals whose noncognitive skills were never measured.

Additionally, several designs leveraging PGS and family relationships allow to disentangle a genetic effect of the offspring's PGS on the offspring's trait (direct genetic effect), and an environmentally-mediated effect of the parental PGS on the offspring's phenotype (parental indirect genetic effect). The presence of indirect genetic effects of noncognitive skills would therefore suggest that environmental factors related to the parents' noncognitive skills also influence the offspring's educational outcomes.

These family-based designs typically leverage parents-offspring trios, siblings (notably twins), or adopted individuals. The idea behind them is either to break up or control for the gene-environment correlation resulting from the transmission of both genes and environment from parents to children. For parents-offspring trios, knowing the genetic variants of both parents and offspring permits to identify genetic variants not transmitted to the child, variants which could therefore only be associated with the child's education via the child's environment. In the case of an adopted child, the child is not genetically related to their adoptive parents, so the estimated effect of the child's PGS on the child trait only reflects their direct genetic effects (no inflation by parental indirect genetic effects). For siblings, comparing the siblings' PGS adjusts for shared genetic effects (which include parental indirect genetic effects). In **Chapter 3**, I extend the adoption and siblings designs to recover an estimate of the parental indirect genetic effects: by contrasting PGS estimates for adopted and non-adopted children and contrasting PGS effects obtained with and without comparing siblings.

To estimate the indirect genetic effects of noncognitive skills on education in **Chapter 3**, I triangulate across these three family-based PGS designs using parents-offspring trio, using adoptees, and using siblings. I do so in three large family cohorts: the Netherlands Twin Register, UK Biobank, and UK Twins Early Development Study. Triangulation across these designs and across cohorts strengthens our conclusions.

As the designs to identify (in)direct genetic effects were new (post 2018), little was known about the sources of biases and the expected direction of biases of each of these approaches. In **Chapter 3**, I additionally performed extensive simulations to investigate the sensitivity of the three designs to common biases in PGS studies. Supplementary sensitivity analyses

were then performed in the three cohorts to estimate the presence of these biases.

Educational attainment and mental health

Previous studies show that the risk of being diagnosed with a mental disorder is higher among those with lower educational attainment[36–38]. **Does this correlation reflect a causal relationship between educational attainment and mental health?** This association could indeed result from the causal effect of education attainment on mental health, or from the causal effect of mental health on educational attainment. Alternatively, it could result from confounding factors influencing both education and mental health, such as childhood socioeconomic status.

Genetically-informed designs can help to examine causality (or its lack thereof). In **Chapter 4**, I apply two quasi-experimental designs that can, given that a set of assumptions are met, test for a plausible causal effect. Triangulating these two designs with different assumptions and bias additionally strengthens our conclusions.

I first employ a within-sibship design: Comparing siblings allows to adjust the association between EA and psychiatric disorders for confounding factors that are shared between siblings, without having to measure them. I contrast this approach with a two-sample mendelian randomization approach. Mendelian randomization[39] leverages the fact that genetic variants can be used as instrumental variables due to two characteristics: genetic variants are fixed at birth and therefore free from reverse causation, and they are randomly inherited from parents. Mendelian randomization, under specific assumptions, therefore provides us with a causal estimate, for a causal effect in a specific direction.

In Dutch national registry data (i.e. Statistics Netherlands, or CBS), I compare the relationship between the number of years of education and diagnoses for 17 psychiatric disorders within 1.7 million siblings. The national register offers an almost unbiased view of DSM-IV diagnosis of 17 psychiatric diagnoses in second-line psychiatric care. For the two-sample mendelian randomization, I use genetic variants associated with EA and 9 psychiatric diagnoses, based on data aggregated across millions of individuals in numerous large GWAS studies.

Parental mental health and children's education

Children of parents with psychopathology generally do less well in school than their peers. However, parental symptoms or disorders need not be responsible for a child's lower achievement, as other familial factors could be at play. As mentioned earlier, genetic transmission or the effect of other confounding factors such as socioeconomic status might create this association. **So does parents' mental health affect the education of their children?**

To account for confounding factors, in **Chapter 5**, I extend the within-sibship design to compare siblings and their children: Is the sibling with worse mental health more likely to have children who do poorly in school? Comparing families whose parents are siblings controls for unmeasured factors shared among adult siblings (e.g. genetics and socioeconomic status) that might confound the relationship between parental mental health and offspring's academic achievement. I analyse data from up to 9,000 families of the longitudinal Norwegian Mother, Father, and Child Study (MoBa). Parents filled out surveys on their symptoms of anxiety, depression, eating disorders, attention deficit hyperactivity disorder, and alcohol problematic use. I specifically focus on the questionnaire administered during pregnancy and at the closest time to the child's outcome. Children in 5th Grade (aged 10) participated in nationally-standardised tests of mathematics, reading comprehension, and English (as an additional language).

Synthesizing findings and future directions

Finally, in **Chapter 6**, I summarize my research. In **Chapter 7**, I integrate the knowledge acquired during this thesis and discuss the potential impact of my research. I reflect on my use of observational data, triangulation, and genetically-informed designs, and suggest future research directions.

The work done in this thesis would not have been possible without the generosity of many. Notably, my analyses relied on the data from the Netherlands Twin Register, the National Longitudinal Study of Adolescent to Adult Health, the Texas Twins Project, the Dunedin Longitudinal Study, the E-Risk Longitudinal Twin Study, the Wisconsin Longitudinal Study, the UK Biobank, the UK Twins Early Development Study, the Norwegian Mother, Father and Child Study, the Dutch and the Norwegian national population registries and many other cohorts whose data was used within published GWAS summary statistics. This dissertation would not have been possible without the participants and everyone else involved with these data.

CHAPTER 2

INVESTIGATING THE GENETIC ARCHITECTURE OF NONCOGNITIVE SKILLS USING GWAS-BY-SUBTRACTION

Published as: **Demange, P.A.***, Malanchini, M.*, Mallard, T.T., Biroli, P., Cox, S.R., Grotzinger, A.D., Tucker-Drob, E.M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D.I., Caspi, A., Corcoran, D.L., Domingue, B. W., Harris, K.M., Ip, H.F., Mitchell, C., Moffitt, T.E., Poulton, R., Prinz, J.A., Sugden, K., Wertz, J., Williams, B.S., de Zeeuw, E.L., Belsky, D.W.*, Harden, K.P.* & Nivard, M.G.* (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nature Genetics*, 53(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Supplementary materials accessible at:

https://www.nature.com/articles/s41588-020-00754-2#Sec32

ABSTRACT

Little is known about the genetic architecture of traits affecting educational attainment other than cognitive ability. We used genomic structural equation modelling and prior genome-wide association studies (GWASs) of educational attainment (N = 1,131,881) and cognitive test performance (N = 257,841) to estimate SNP associations with educational attainment variation that is independent of cognitive ability. We identified 157 genome-wide significant loci and a polygenic architecture accounting for 57% of genetic variance in educational attainment. Noncognitive genetics were enriched in the same brain tissues and cell types as cognitive performance but showed different associations with grey-matter brain volumes. Noncognitive genetics were further distinguished by associations with personality traits, less risky behaviour, and increased risk for certain psychiatric disorders. For socioeconomic success and longevity, noncognitive and cognitive-performance genetics demonstrated associations of similar magnitude. By conducting a GWAS of a phenotype that was not directly measured, we offer a view of genetic architecture of noncognitive skills influencing educational success.

INTRODUCTION

"It takes something more than intelligence to act intelligently."
– Fyodor Dostoyevsky, *Crime and Punishment*

Success in school – and life – depends on skills beyond cognitive ability. Randomized trials of early life education interventions find substantial benefits to educational outcomes, employment and adult health, even though the interventions have no lasting effects on children's cognitive functions[40,41]. These results have captured the attention of educators and policy-makers, motivating interest in so-called "noncognitive skills"[42–44]. Noncognitive skills suspected to be important for educational success include motivation, curiosity, persistence, and self-control[26–28,45,46]. However, questions have been raised about the substance of these skills and the magnitudes of their impacts on life outcomes[47].

Twin studies find evidence that noncognitive skills are heritable[48–52]. Genetic analysis could help clarify the contribution of these skills to educational attainment and elucidate their connections with other traits. However, lack of consistent and reliable measurements of noncognitive skills in existing genetic datasets poses challenges[30].

To overcome these challenges, we designed a GWAS of a latent trait, that is, a trait not measured in any of the genotyped subjects[53]. We borrowed the strategy used in the original analysis of noncognitive skills within the discipline of economics[54,55]: we defined genetic influences on noncognitive skills as the genetic variation in educational attainment that was not explained by cognitive skills. We then performed GWAS on this residual "noncognitive" genetic variation in educational attainment. This approach is a necessarily imperfect representation of the true relationship between cognitive and noncognitive skills; in human development, cognitive abilities and other skills relevant for educational attainment probably interact dynamically, each influencing the other[56]. Our analysis excludes genetic influences on education-relevant skills that also influence measured cognitive abilities. The value of this imperfect approach is to make a quantity otherwise difficult to study tractable for analysis.

We conducted analysis using Genomic Structural Equation Modelling (Genomic-SEM)[57] applied to published GWAS summary statistics for educational attainment and cognitive performance[58]. Our analysis used these summary statistics to "subtract" genetic influence on cognitive performance from the association of each SNP with educational attainment. The remaining associations of each SNP with educational attainment formed a new GWAS of a noncognitive skills phenotype that was never directly measured. We call this new statistical approach GWAS-by-subtraction.

We used results from the GWAS-by-subtraction of noncognitive skills to conduct two sets of analyses. First, we conducted hypothesis-driven analysis using the phenotypic annotation approach[59]. We used genetic correlation and polygenic score analysis to test the hypothesis that noncognitive skills influence educational and economic attainments and longevity and to investigate traits and behaviours that constitute noncognitive skills. Second, we conducted hypothesis-free bioinformatic annotation analysis to explore the tissues, cell-types, and brain structures that might distinguish the biology of noncognitive skills from the biology mediating cognitive influences on educational attainment.

RESULTS

GWAS-by-subtraction identifies genetic associations with noncognitive variance in educational attainment.

The term "noncognitive skills" was originally coined by economists studying individuals who were equivalent in cognitive ability but differed in educational attainment[55]. Our analysis of noncognitive skills was designed to mirror this original approach: we focused on genetic variation in educational outcomes not explained by genetic variation in cognitive ability. Specifically, we applied Genomic-SEM[57] to summary statistics from GWASs of educational attainment[58] and cognitive performance[58]. Both phenotypes were regressed on a latent factor representing genetic variance in cognitive performance (hereafter "Cog"). Educational attainment was further regressed on a second latent factor representing the residual genetic variance in educational attainment left over after regressing out variance related to cognitive performance (hereafter "NonCog"). By construction, NonCog genetic variance was independent of Cog genetic variance ($r_g = 0$). In other words, the NonCog factor represents genetic variation in educational attainment that is not accounted for by the Cog factor. These two latent factors were then regressed on individual SNPs, yielding a GWAS of the latent constructs NonCog and Cog. A graphical representation of the model is presented in Figure 1. Parameters are derived in terms of the observed moments of the joint distribution of educational attainment, cognitive performance, and an SNP (see Supplementary Note).

The NonCog latent factor accounted for 57% of total genetic variance in educational attainment. Using linkage disequilibrium (LD) score regression[60], we estimated SNP heritability for NonCog to be $h^2_{NonCog} = 0.0637$ ($SE = 0.0021$). After conventional GWAS significance threshold correction, GWAS of NonCog identified 157 independent genome-wide-significant lead SNPs (independent SNPs defined as outside a 250-kb window, or within a 250-kb window and $r^2 < 0.1$). The results from the NonCog GWAS are shown as a Manhattan plot in Figure 2. NonCog and Cog GWAS details are reported in Supplementary

Tables 1-4, Supplementary Figure 1, and the Supplementary Note. In addition, we report a series of sensitivity analyses as follows: analysis of potential biases due to cohort differences (Supplementary Table 5 and Supplementary Figures 2-4); analysis of impact of allowing for positive genetic correlations between NonCog and Cog (Supplementary Tables 6 and 7, and Supplementary Figures 5 and 6; analysis of impact of allowing for a moderate causal effect of educational attainment on cognitive performance[24] (Supplementary Table 8 and Supplementary Figures 7-9).

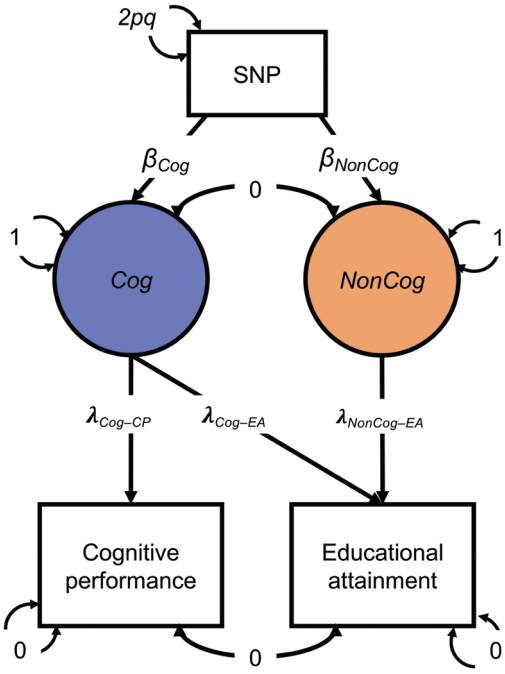


Figure 1 | GWAS-by-subtraction Genomic-SEM model. Cholesky model as fitted in Genomic-SEM, with path estimates for a single SNP included as illustration. SNP, cognitive performance (CP), and educational attainment (EA) are observed variables based on GWAS summary statistics. The genetic covariance between CP and EA is estimated based on their GWAS summary statistics. The model is fitted to a 3 x 3 observed variance-covariance matrix (i.e. SNP, CP, EA). Cog and NonCog are latent (unobserved) variables. The covariances between CP and EA and between Cog and NonCog are fixed to 0. The variance of the SNP is fixed to the value of $2pq$ (p = reference allele frequency, q = alternative allele frequency, based on 1000 Genomes phase 3). The residual variances of CP and EA are fixed to 0, so that all variance is explained by the latent factors. The variances of the latent factors are fixed to 1. The observed variables CP and EA were regressed on the latent variables resulting in the estimates for the path loadings: $\lambda_{Cog\text{-}CP} = 0.4465$; $\lambda_{Cog\text{-}EA} = 0.2237$; $\lambda_{NonCog\text{-}EA} = 0.2565$. The latent variables were then regressed on each SNP that met QC criteria.

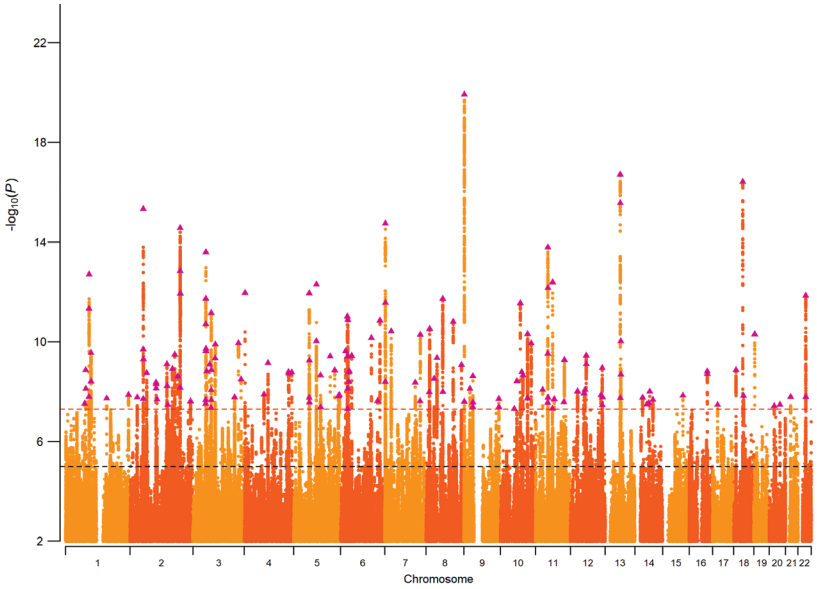


Figure 2 | Manhattan plot of SNP associations with NonCog. Plot of the -$\log_{10}$(p-value) associated with the Wald test (two-sided) of β_{NonCog} for all SNPs, ordered by chromosome and base position. Purple triangles indicate genome-wide significant ($p < 5 \times 10^{-8}$) and independent (within a 250-kb window and $r^2 < 0.1$) associations. The red dashed line marks the threshold for genome-wide significance ($p = 5 \times 10^{-8}$), and the black dashed line the threshold for nominal significance ($p = 1 \times 10^{-5}$).

Phenotypic annotation analysis elucidates correlates of noncognitive skills genetics.

Our phenotypic annotation analyses proceeded in two steps. First, we conducted polygenic score (PGS) and genetic correlation (r_g) analysis to test whether our GWAS-by-subtraction succeeded in identifying genetic influences that were important to educational attainment and also distinct from genetic influences on cognitive ability. Second, we conducted PGS and r_g analyses to explore how NonCog related to a network of phenotypes that psychology and economics research suggests might form the basis of noncognitive influences on educational attainment.

NonCog genetics are associated with education, socioeconomic attainment and longevity.

To establish whether the Genomic-SEM GWAS-by-subtraction succeeded in isolating genetic variance in education that was independent of cognitive function, we compared genetic associations of NonCog and Cog with educational attainment and cognitive test performance. Results for analysis of education and cognitive test phenotypes are shown in Figure 3.

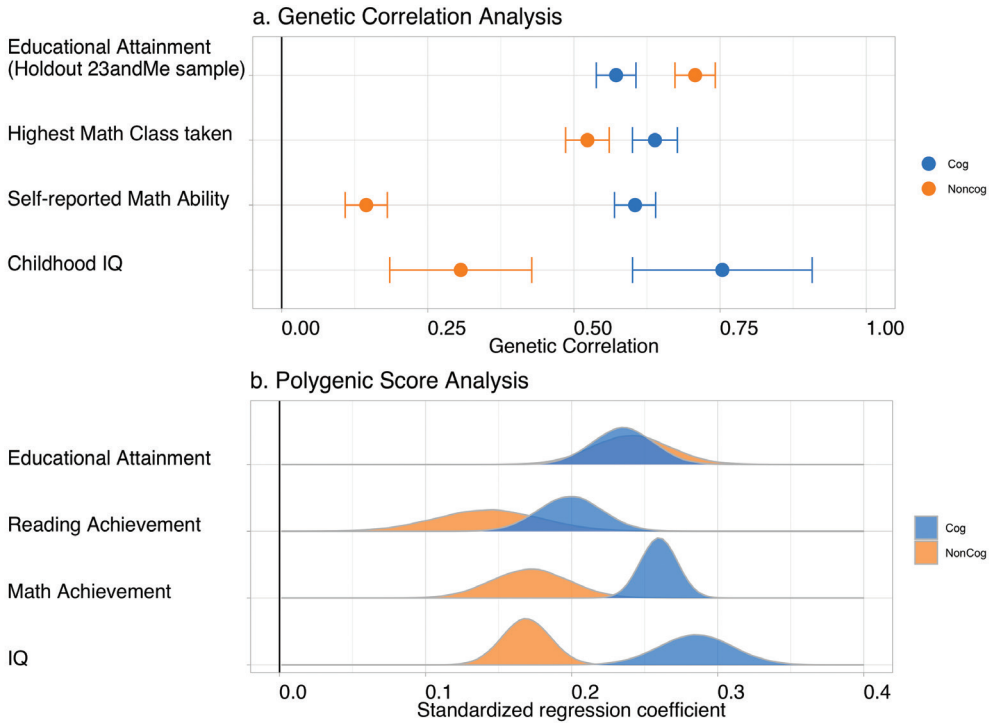


Figure 3 | Polygenic prediction and genetic correlations with IQ and educational achievement. **a,** Genetic correlations of NonCog and Cog with educational attainment, highest math class taken, self-reported math ability, and childhood IQ. The dots represent genetic correlations estimated using Genomic-SEM. Correlations with NonCog are in orange, and with Cog in blue. Error bars represent 95% CIs. Exact estimates and P-values are reported in Supplementary Table 14. For analysis of genetic correlations with educational attainment, we re-ran the Genomic-SEM model to compute NonCog and Cog using summary statistics that omitted the 23andMe sample from the educational attainment GWAS. We then used the 23andMe sample to run the GWAS of educational attainment. Thus, there is no sample overlap in this analysis. **b,** Effect-size distributions from meta-analysis of NonCog and Cog polygenic score associations with cognitive test performance and educational attainment. Outcomes were regressed simultaneously on NonCog and Cog polygenic scores. Effect-sizes entered into the meta-analysis were standardized regression coefficients interpretable as Pearson r. Exact estimates and P-values are reported in Supplementary Table 12. Samples and measures are detailed in Supplementary Tables 9 and 10. Traits were measured in different samples: educational attainment was measured in the AddHealth, Dunedin, E-Risk, NTR and WLS samples (N = 24,056); reading achievement and mathematics achievement were measured in the AddHealth, NTR, and Texas-Twin samples (N = 9,274 for reading achievement; N = 10,747 for mathematics achievement); cognitive test performance (IQ) was measured in the Dunedin, E-Risk, NTR, Texas Twins and WLS samples (N = 11,351). The densities were obtained by randomly generating normal distributions where the meta-analytic estimate was included as the mean and the meta-analytic standard error as the standard deviation.

We conducted PGS analysis of educational attainment in the Netherlands Twin Register[61] (NTR), National Longitudinal Study of Adolescent to Adult Health[62] (AddHealth), Dunedin Longitudinal Study[63], E-Risk6[4], and Wisconsin Longitudinal Study[65] (WLS) cohorts (meta-analysis N = 24,056; cohort descriptions in Supplementary Tables 9 and 10 and Supplementary Note). PGS effect sizes were the same for NonCog and Cog (NonCog β = 0.24 (*SE* = 0.03), Cog β = 0.24 (*SE* = 0.02), p_{diff} = 0.702; all PGS results are reported in Supplementary Tables 11 and 12). We conducted complementary genetic correlation analysis using Genomic-SEM and GWAS summary statistics from a hold-out-sample GWAS of educational attainment (Supplementary Note). This analysis allowed us to compute an out-of-sample genetic correlation of NonCog with educational attainment. NonCog showed a stronger genetic correlation with educational attainment as compared to Cog (NonCog r_g = 0.71 (*SE* = 0.02), Cog r_g = 0.57 (*SE* = 0.02), p_{diff} < 0.0001; all genetic correlation results are reported in Supplementary Tables 13 and 14).

We conducted PGS analysis of cognitive test performance in the NTR, Texas Twin Project[66], Dunedin, E-Risk, and WLS cohorts (combined N = 11,351). The goal of our GWAS-by-subtraction analysis was to exclude, as much as possible, genetic variance in cognitive ability from genetic variance in skills relevant for education. Consistent with this goal, effect-sizes for NonCog PGS associations with full-scale intelligence quotient (IQ) were smaller by half as compared to Cog PGS associations (NonCog β = 0.17 (*SE* = 0.02), Cog β = 0.29 (*SE* = 0.03); p_{diff} < 0.0001). However, the non-zero correlation between the NonCog PGS and full-scale IQ is a reminder that the cognitive performance GWAS used in our GWAS-by-subtraction analyses does not capture the entirety of genetic influences on all forms of cognitive tests measured at all points in the lifespan. Additional PGS analyses of IQ subscales are reported in Supplementary Figure 10 and Supplementary Tables 11 and 12.

We conducted complementary genetic correlation analysis using results from a published GWAS of childhood IQ[67]. Parallel to PGS analysis, the NonCog genetic correlation with childhood IQ was smaller by more than half as compared to the Cog genetic correlation (NonCog r_g = 0.31 (*SE* = 0.06), Cog r_g = 0.75 (*SE* = 0.08), p_{diff_fdr} < 0.0001). Of the total genetic correlation between childhood IQ and educational attainment, 31% of the covariance was explained by NonCog and 69% by Cog.

We next examined downstream economic and health outcomes associated with greater educational attainment[68,69]. In PGS analysis in the AddHealth and Dunedin cohorts (N = 6,358), NonCog and Cog PGSs showed similar associations with occupational attainment (NonCog β = 0.21 (*SE* = 0.01), Cog β = 0.21 (*SE* = 0.01), p_{diff} = 0.902). In genetic correlation analysis, NonCog showed a similar relationship to income[70] as Cog (NonCog r_g = 0.62, (*SE* =

0.04), Cog r_g = 0.62 (*SE* = 0.04), p_{diff_fdr} = 0.947) and a stronger relationship with neighborhood deprivation[70], a measure related to where a person can afford to live (NonCog r_g = -0.51 (*SE* = 0.05), Cog r_g = -0.32 (*SE* = 0.04), p_{diff_fdr} = 0.001). In Genomic-SEM analysis, NonCog explained 53% of the genetic correlation between educational attainment and income and 65% of the genetic correlation between educational attainment and neighborhood deprivation (Supplementary Table 15).

We conducted genetic correlation analysis of longevity based on GWAS of parental lifespan[71]. Genetic correlations were stronger for NonCog as compared to Cog (NonCog r_g = 0.37 (*SE* = 0.03); Cog r_g = 0.27 (*SE* = 0.03); p_{diff_fdr} = 0.024). In Genomic-SEM analysis, NonCog explained 61% of the genetic correlation between educational attainment and longevity.

In summary, NonCog and Cog genetics showed similar relationships with educational attainment and its long-term outcomes, despite NonCog genetic having a much weaker relationship to measured cognitive test performance than Cog genetics. These findings broadly support the hypothesis that noncognitive skills distinct from cognitive abilities are an important contributor to success across the life course.

We next conducted a series of genetic correlation analyses to explore the network of phenotypes to which NonCog was genetically correlated. To develop understanding of the substance of noncognitive skills, we tested where in that network of phenotypes genetic correlations with NonCog diverged from genetic correlations with Cog. Our analysis was organized around four themes: decision-making preferences, health-risk and fertility behaviours, personality traits, and psychiatric disorders. Results of genetic correlation analyses are shown in Figure 4 and Supplementary Figure 11. Results are reported in Supplementary Table 14.

NonCog genetics were associated with decision-making preferences.

In economics, noncognitive influences on achievement and health are often studied in relation to decision-making preferences[72–75]. NonCog was genetically correlated with higher tolerance of risks[76] (r_g = 0.10 (*SE* = 0.03)) and willingness to forego immediate gratification in favor of a larger reward at a later time[77] (delay discounting r_g = -0.52 (*SE* = 0.08)). In contrast, Cog was genetically correlated with generally more cautious decision-making characterized by lower levels of risk tolerance (r_g = -0.35 (*SE* = 0.07), p_{diff_fdr} < 0.0001) and delay discounting (r_g = -0.35 (*SE* = 0.07), p_{diff_fdr} = 0.082).

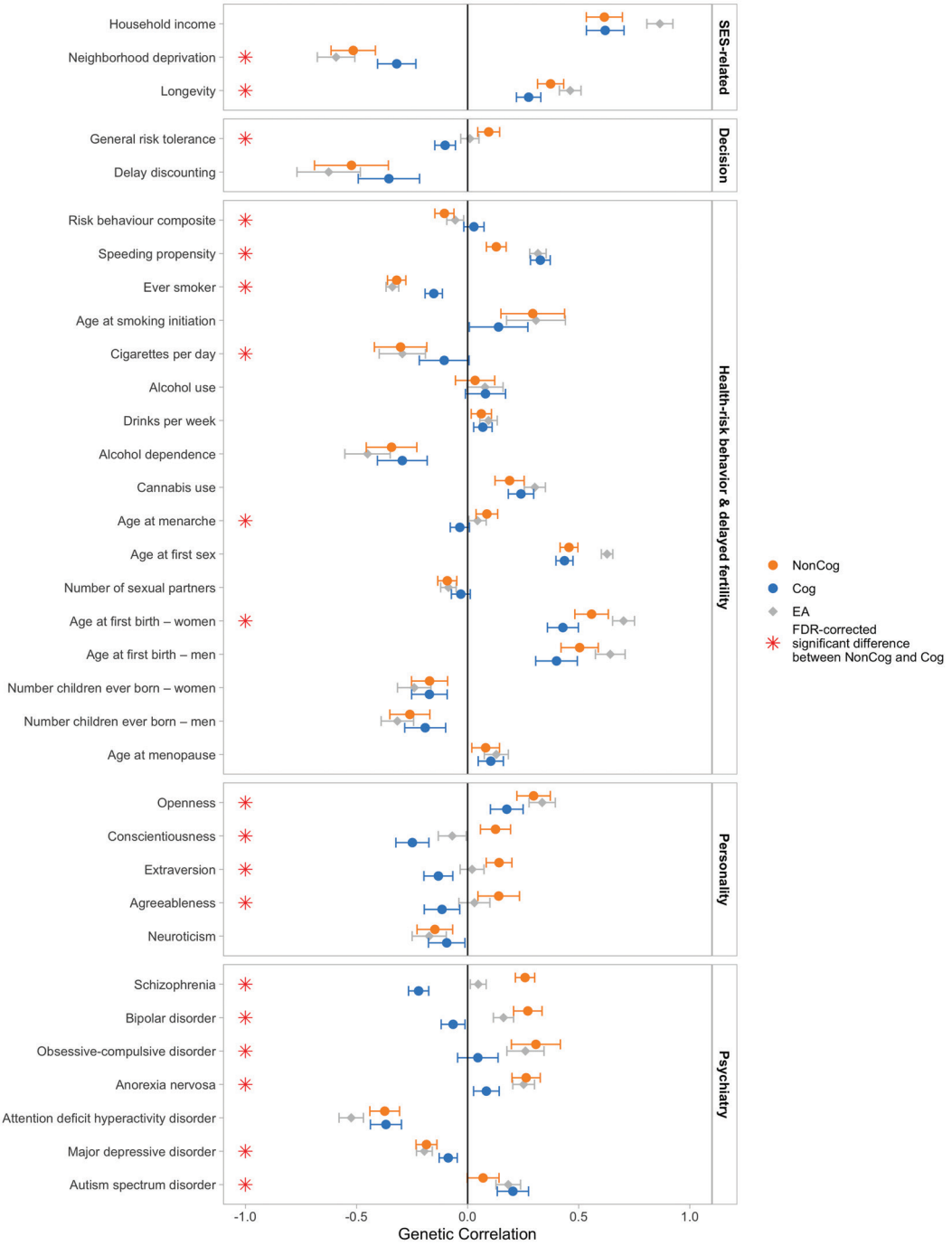

Household income
Neighborhood deprivation
Longevity
SES-related
General risk tolerance
Delay discounting
Decision
Risk behaviour composite
Speeding propensity
Ever smoker
Age at smoking initiation
Cigarettes per day
Alcohol use
Drinks per week
Alcohol dependence
Cannabis use
Age at menarche
Age at first sex
Number of sexual partners
Age at first birth – women
Age at first birth – men
Number children ever born – women
Number children ever born – men
Age at menopause
Health-risk behavior & delayed fertility
Openness
Conscientiousness
Extraversion
Agreeableness
Neuroticism
Personality
Schizophrenia
Bipolar disorder
Obsessive-compulsive disorder
Anorexia nervosa
Attention deficit hyperactivity disorder
Major depressive disorder
Autism spectrum disorder
Psychiatry
-1.0
-0.5
0.0
0.5
1.0
Genetic Correlation
NonCog
Cog
EA
FDR-corrected significant difference between NonCog and Cog

Figure 4 | Estimates of genetic correlations with NonCog, Cog, and educational attainment. Genetic correlations of NonCog, Cog, and educational attainment with selected phenotypes. The dots represent genetic correlations estimated in Genomic-SEM. Correlations with NonCog are in orange, with Cog in blue, and with educational attainment in gray. Error bars represent 95% CIs. Red stars indicate a statistically significant (FDR corrected $p < 0.05$, two-tailed test) difference in the magnitude of the correlation with NonCog versus Cog. Exact *p*-values for all associations are reported in Supplementary Table 14. The FDR correction was applied based on all genetic correlations tested (including in Supplementary Figure 11). The difference test is based on a chi-squared test associated with a comparison between a model constraining these two correlations to be identical versus a model where the correlations are freely estimated. Source GWASs are listed in Supplementary Table 13.

NonCog genetics were associated with less health-risk behavior and delayed fertility.

An alternative approach to studying specific noncognitive skills is to infer individual differences in noncognitive skills from patterns of health-risk behavior. NonCog was genetically correlated with less health-risk behavior as indicated by analysis of obesity[78], substance use[76,79–82], and sexual behaviors and early fertility[76,83,84] (r_g range 0.2–0.5), with the exception that the r_g with alcohol use was not different from zero and r_g with cannabis use was positive. Genetic correlations for Cog were generally in the same direction but of smaller magnitude.

NonCog genetics were associated with personality characteristics linked with social and professional competency.

In psychology, noncognitive influences on achievement are conceptualized as personality traits, *i.e.* patterns of stable individual differences in emotion and behavior. The model of personality that has received the most attention in genetics is a five-factor model referred to as the Big Five. Genetic correlation analysis of the Big Five personality traits[85–87] revealed that NonCog genetics were most strongly associated with Openness to Experience (being curious and eager to learn; $r_g = 0.30$ ($SE = 0.04$)) and were further associated with a pattern of personality characteristic of changes that occur as people mature in adulthood[88]. Specifically, NonCog showed a positive r_g with Conscientiousness (being industrious and orderly; $r_g = 0.13$ ($SE = 0.03$)), Extraversion (being enthusiastic and assertive; $r_g = 0.14$ ($SE = 0.03$)), and Agreeableness (being polite and compassionate; $r_g = 0.14$ ($SE = 0.05$)), and negative r_g with Neuroticism (being emotionally volatile; $r_g = -0.15$ ($SE = 0.04$)). Genetic correlations of Cog with Openness to Experience and Neuroticism were similar to those for NonCog ($p_{diff\_fdr\text{-}Openness} = 0.040$, $p_{diff\_fdr\text{-}Neuroticism} = 0.470$). In contrast, genetic correlations of Cog with Conscientiousness, Extraversion, and Agreeableness were in the opposite direction ($r_g = -0.25$ to -0.12, $p_{diff\_fdr} < 0.0005$). PGS analysis of personality traits is reported in Supplementary Table 12, Supplementary Figure 12, and the Supplementary Note.

NonCog genetics were associated with higher risk for multiple psychiatric disorders.

In clinical psychology and psychiatry, research is focused on mental disorders. Mental disorders are generally associated with impairments in academic achievement and social role functioning[89,90]. However, positive genetic correlations with educational attainment and creativity have been reported for some disorders[91,92]. We therefore tested NonCog r_g with psychiatric disorders based on published case-control GWASs of mental disorders[93–99]. NonCog was associated with higher risk for multiple clinically defined disorders, including anorexia nervosa (r_g = 0.26 (*SE* = 0.04)), obsessive-compulsive disorder (r_g = 0.31 (*SE* = 0.06)), bipolar disorder (r_g = 0.27 (*SE* = 0.03)), and schizophrenia (r_g = 0.26 (*SE* = 0.02)). Genetic correlations between Cog and psychiatric disorders were either smaller in magnitude (anorexia nervosa r_g = 0.08 (*SE* = 0.03), $p_{diff\_fdr} < 0.001$; obsessive-compulsive disorder r_g = 0.05 (*SE* = 0.05), p_{diff_fdr} = 0.002) or in the opposite direction (bipolar disorder r_g = -0.07 (*SE* = 0.03), $p_{diff\_fdr} < 0.001$; schizophrenia r_g = -0.22 (*SE* = 0.02), $p_{diff\_fdr} < 0.001$). Both NonCog and Cog showed negative genetic correlations with attention deficit hyperactivity disorder (NonCog r_g = -0.37 (*SE* = 0.03), Cog r_g = -0.37 (*SE* = 0.04), p_{diff_fdr} = 0.947).

In summary, NonCog genetics were associated with phenotypes from economics and psychology thought to mediate noncognitive influences on educational success. These associations contrasted with associations for Cog genetics, supporting distinct pathways of influence on achievement in school and later in life. Opposing patterns of association were also observed for psychiatric disorders, suggesting that the unexpected positive genetic correlation between educational attainment and mental health problems uncovered in previous studies[92,100,101] arises from noncognitive genetic influences on educational attainment.

Biological annotation analyses reveal shared and specific neurobiological correlates.

The goal of biological annotation of GWAS discoveries is to elucidate molecular mechanisms mediating genetic influences on the phenotype of interest. Our biological annotation analysis proceeded in two steps. First, we conducted enrichment analysis to test whether some tissues and cell types were more likely to mediate NonCog and Cog heritabilities than others. Second, we conducted genetic correlation analysis to explore how NonCog and Cog genetics related to different brain structures.

NonCog and Cog genetics were enriched in similar tissues and cells.

We tested whether common variants in genes specifically expressed in 53 Genotype-

Tissue Expression (GTEx) tissues[102] or in 152 tissues captured in a previous aggregation of RNA-sequencing studies[103,104] were enriched in their effects on Cog or NonCog. Genes predominantly expressed in the brain rather than peripheral tissues were enriched in both NonCog and Cog (Supplementary Table 16).

To examine expression patterns at a more granular level of analysis, we used MAGMA[105] and stratified the LD score regression[106] to test enrichment of common variants in 265 nervous system cell-type-specific gene sets[107] (Supplementary Table 17). In MAGMA analysis, common variants in 95 of 265 gene-sets were enriched for association with NonCog. The enriched cell types were predominantly neurons (97%), with enrichment most pronounced for telencephalon-projecting neurons, di- and mesencephalon neurons, and to a lesser extent, telencephalon interneurons (Supplementary Figure 13 and Supplementary Table 18). Enrichment for Cog was similar to NonCog (correlation between Z-statistics Pearson's r = 0.85), and there were no differences in cell-type-specific enrichment, suggesting that the same types of brain cells mediate genetic influences on NonCog and Cog (Supplementary Figure 14). Stratified LDSC results were similar to results from MAGMA (Supplementary Note, Supplementary Figure 15, and Supplementary Table 19).

The absence of differences in cell-type-specific enrichment is surprising given that NonCog and Cog are genetically uncorrelated. We therefore used the TWAS/Fusion tool[108] to conduct gene-level analysis. This analysis revealed a mixture of concordant and discordant gene effects on NonCog and Cog consistent with the genetic correlation of 0 (Supplementary Note, Supplementary Figure 16, and Supplementary Table 20).

NonCog and Cog genetics show diverging associations with total and regional brain volumes.

Educational attainment has previously been found to be genetically correlated with greater total brain volume[109,110]. We therefore used a GWAS of regional brain volume to compare the r_g of NonCog and Cog with total brain volume and 100 regional brain volumes (99 grey-matter volumes and 1 white matter volume) controlling for total brain volume (Supplementary Table 21)[111]. For total brain volume, genetic correlation was stronger for Cog as compared to NonCog (Cog r_g = 0.22 (SE = 0.04), NonCog r_g = 0.07 (SE = 0.03), p_{diff} = 0.005). Total grey-matter volume, controlling for total brain volume, was not associated with either NonCog or Cog (NonCog: r_g = 0.07 (SE = 0.04); Cog: r_g = 0.06 (SE = 0.04)). For total white matter volume, conditional on total brain volume, genetic correlation was weakly negative for NonCog compared with Cog (NonCog r_g = -0.12 (SE = 0.04), Cog (r_g = -0.01 (SE = 0.04), p_{diff} = 0.04).

NonCog was not associated with any of the regional grey-matter volumes after false discovery rate (FDR) correction. In contrast, Cog was significantly associated with regional

grey-matter volumes for the bilateral fusiform, insula and posterior cingulate (r_g range 0.11-0.17), as well as left superior temporal (r_g = 0.11 (*SE* = 0.04)), left pericalcarine (r_g = -0.16 (*SE* = 0.05)) and right superior parietal volumes (r_g = -0.22 (*SE* = 0.06)) (Figure 5).

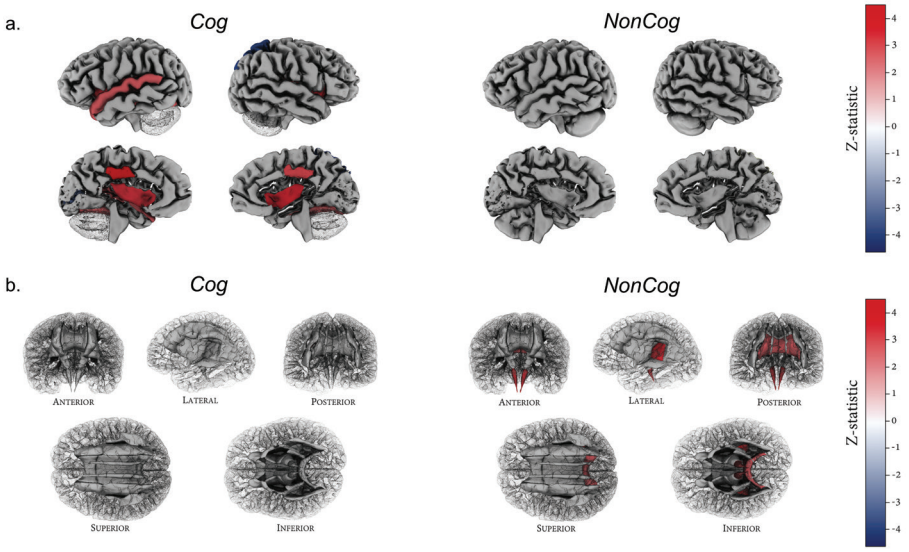


Figure 5 | Genetic correlations with regional gray-matter volumes and white-matter tracts. a. Cortical patterning of FDR-corrected significant genetic correlations with regional gray-matter volumes for Cog versus NonCog, after correction for total brain volume. Regions of interest are plotted according to the Desikan-Killiany-Tourville atlas[112], shown on a single manually-edited surface (http://mindboggle.info[113]). Exact estimates and *p*-values are reported in Supplementary Table 21. Cog showed significant associations with gray-matter volume for the bilateral fusiform, insula and posterior cingulate, the left superior temporal and left pericalcarine and right superior parietal volumes. NonCog was not associated with any of the regional brain volumes. **b.** White matter tract patterning of FDR-corrected significant genetic correlations with regional mode of anisotropy (MO) for Cog versus NonCog. White matter tract probability maps are plotted according to the Johns Hopkins University DTI atlas (https://identifiers.org/neurovault.image:1401)[114]. Exact estimates and *p*-values are reported in Supplementary Table 21. Cog was not associated with regional MO. NonCog showed significant associations with MO in the corticospinal tract, the retrolenticular limb of the internal capsule and the splenium of the corpus callosum.

Finally, we tested genetic correlation of NonCog and Cog with white-matter tract integrity as measured using diffusion tensor imaging (DTI)[115]. Analyses included 5 DTI parameters in each of 22 white-matter tracts (Supplementary Table 22). NonCog was positively associated with the mode of anisotropy parameter (which denotes a more tubular, as opposed to planar, water diffusion) in the corticospinal tract, retrolenticular limb of the internal capsule, and splenium of the corpus callosum (Figure 5). However, all correlations were small ($0.10 < r_g < 0.14$), and we detected no genetic correlations that differed between NonCog and Cog (Supplementary Note).

DISCUSSION

GWAS of noncognitive influences on educational attainment identified 157 independent loci and polygenic architecture accounting for more than half the genetic variance in educational attainment. In genetic correlation and PGS analysis, these NonCog genetics showed a similar magnitude of associations with educational attainment, economic attainment, and longevity to genetics associated with cognitive influences on educational attainment (Cog). As expected, NonCog genetics had much weaker associations with cognition phenotypes compared with Cog genetics. These results contribute new GWAS evidence in support of the hypothesis that heritable noncognitive skills influence educational attainment and downstream life-course economic and health outcomes.

Phenotypic and biological annotation analyses shed light on the substance of heritable noncognitive skills influencing education. Economists hypothesize that preferences that guide decision-making in the face of risk and delayed rewards represent noncognitive influences on educational attainment. Consistent with this hypothesis, NonCog genetics were associated with higher risk tolerance and lower time discounting. These decision-making preferences are associated with financial wealth, whereas the opposite preferences are hypothesized to contribute to a feedback loop perpetuating poverty[116]. Consistent with results from the analysis of decision-making preferences, NonCog genetics were also associated with healthier behaviour and later fertility.

Psychologists hypothesize that the Big Five personality characteristics of conscientiousness and openness are the two "pillars of educational success"[48,117,118]. Our results provide some support for this hypothesis, with the strongest genetic correlation evident for openness. However, they also show that noncognitive skills encompass the full range of personality traits, including agreeableness, extraversion and the absence of neuroticism. This pattern mirrors the pattern of personality change that occurs as young people mature into adulthood[88]. Thus, noncognitive skills share genetic aetiology with what might be termed as "mature personality". The absolute magnitudes of genetic correlations between NonCog and individual personality traits are modest. This result suggests that the personality traits described by psychologists capture some, but not all, genetic influence on noncognitive skills.

Although the general pattern of findings in our phenotypic annotation analysis indicated noncognitive skills were genetically related to socially desirable characteristics and behaviours, there was an important exception. Genetic correlation analysis of psychiatric disorder GWASs revealed positive associations of NonCog genetics with schizophrenia, bipolar disorder, anorexia nervosa, and obsessive-compulsive disorder. Previously, these

psychiatric disorders have been shown to have a positive r_g with educational attainment, a result that has been characterized as paradoxical given the impairments in educational and occupational functioning typical of serious mental illness. Our results clarify that these associations are driven by noncognitive factors associated with success in education. These results align with the theory that clinically defined psychiatric disorders represent extreme manifestations of dimensional psychological traits, which might be associated with adaptive functioning within the normal range[119–121].

Finally, biological annotation analyses suggested that genetic variants contributing to educational attainment not mediated through cognitive abilities are enriched in genes expressed in the brain, specifically in neurons. Even though NonCog and Cog were genetically uncorrelated, variants in the same neuron-specific gene-sets were enriched for both traits. Although we found some evidence of differences between NonCog and Cog in associations with grey matter volumes, moderate sample sizes in neuroimaging GWASs mean these results must be treated as preliminary, requiring replication with data from larger-scale GWASs of white-matter and grey-matter phenotypes. Limited differentiation of NonCog and Cog in biological annotation analyses focused at the levels of tissue and cell type highlights need for finer-grained molecular data resources to inform these analyses and the complementary value of phenotypic annotation analyses focused at the level of psychology and behaviour.

We acknowledge limitations. Cognitive and noncognitive skills develop in interaction with one another. For example, the dynamic mutualism hypothesis[122] proposes that noncognitive characteristics shape investments of time and effort, leading to differences in the pace of cognitive development[123,124]. However, in Genomic-SEM analysis, the NonCog factor is, by construction, uncorrelated with genetic influences on adult cognition as measured in the Cog GWAS. Our statistical separation of NonCog from cognition is thus a simplified representation of development. Longitudinal studies with repeated measures of cognitive and candidate noncognitive skills are needed to study their reciprocal relationships across development[125,126]. Our statistical separation of NonCog from cognition is also incomplete. The ability to control statistically for any variable, genetic or otherwise, depends on how well and comprehensively that variable is measured[127]. The tests of cognitive performance included in the Cog GWAS probably do not capture all genetic influences on all forms of cognitive ability across the lifespan[128,129]. Despite these limitations, our simplified and incomplete statistical separation of NonCog from Cog allowed us to test whether heritable traits other than cognitive ability influenced educational attainment and to explore what those traits might be.

As our analysis was based on GWAS of educational attainment, noncognitive genetics identified in the present study may differ from noncognitive genetics affecting other socioeconomic attainments like income, or traits and behaviours that mediate responses to early childhood interventions, to the extent that those genetics do not affect educational attainment. Parallel analysis of alternative attainment phenotypes will clarify the specificity of discovered noncognitive genetics.

In the case of the GWAS of educational attainment, the included samples were drawn mainly from western Europe and the USA, and participants completed their education in the late 20th and early 21st centuries. The phenotype of educational attainment reflects an interaction between an individual and the social system in which they are educated. Differences across social systems, including education policy, culture, and historical context, may result in different heritable traits influencing educational attainment[130]. Results therefore may not generalize beyond the times and places GWAS samples were collected.

Generalization of the NonCog factor is also limited by restriction of the included GWASs to individuals of European ancestry. Lack of methods for integrating genome-scale genetic data across populations with different ancestries[131,132] requires this restriction, but raises threats to external validity. GWASs of other ancestries and development of methods for trans-ancestry analysis can enable analysis of (Non)Cog in non-European populations.

Within the bounds of these limitations, results illustrate the application of Genomic-SEM to conduct GWASs of a phenotype not directly measured in GWAS databases. This application could have broad utility beyond the genetics of educational attainment. The GWAS-by-subtraction method allowed us to study a previously hard-to-interpret residual value. Our analysis provides a view of the genetic architecture of noncognitive skills influencing educational success. These skills are central to theories of human capital formation within the social and behavioural sciences and are increasingly the targets of social policy interventions. Our results establish that noncognitive skills are central to the heritability of educational attainment and illuminate connections between genetic influences on these skills and social and behavioural science phenotypes.

METHODS

Meta-analysis of educational attainment GWAS

We reproduced the Social Science Genetic Association Consortium (SSGAC) 2018 GWAS of educational attainment[58] by meta-analysing published summary statistics for N = 766,345 (www.thessgac.org/data) with summary statistics obtained from 23andMe, Inc. (N = 365,538). We included SNPs with sample size > 500,000 and minor allele frequency

> 0.005 in the 1000 Genomes Project reference set (10,101,243 SNPs). We did not apply genomic control, as standard errors of publicly available and 23andMe summary statistics were already corrected[58]. Meta-analysis was performed using METAL[133].

GWAS-by-subtraction

The objective of our GWAS-by-subtraction analysis was to estimate, for each SNP, the association with educational attainment that was independent of that SNP's association with cognition (hereafter, the NonCog SNP effect). We used Genomic-SEM[57] in R 3.4.3 to analyse GWAS summary statistics for the educational attainment and cognitive performance phenotypes in the SSGAC's 2018 GWAS[58]. The model regressed the educational attainment and cognitive performance summary statistics on two latent variables, Cog and NonCog (Figure 1). Cog and NonCog were then regressed on each SNP in the genome. This analysis allowed for two paths of association with educational attainment for each SNP. One path was fully mediated by Cog. The other path was independent of Cog and measured the noncognitive SNP effect, NonCog. To identify independent hits with $p < 5 \times 10^{-8}$ (the customary p-value threshold to approximate an alpha value of 0.05 in GWAS), we pruned the results using a radius of 250 kb and an LD threshold of $r^2 < 0.1$ (Supplementary Tables 1-3). We explore alternative lead SNPs and loci definition in Supplementary Table 4. The parameters estimated in a GWAS-by-subtraction and their derivation in terms of the genetic covariance are described in the Supplementary Note (model specification), and practical analysis steps are further described in the Supplementary Note (SNP filtering). The effective sample size of the NonCog and Cog GWAS was estimated to 510,795 and 257,700, respectively (Supplementary Note). We investigated biases from unaccounted-for heterogeneity in overlap across SNPs in the educational attainment and cognitive performance GWASs and describe a possible strategy to deal with it (Supplementary Note). We investigated potential biases due to cohort differences in SNP heritability in the Supplementary Note. We evaluated the consequences of modifying r_g(NonCog, Cog) = 0 by evaluating r_g = 0.1, 0.2 or 0.3, and we investigated the consequences of a violation of the assumed causation between cognitive performance and educational attainment in the Supplementary Note.

PGS analysis

PGS analyses were conducted in data drawn from six population-based cohorts from the Netherlands, the UK, the USA, and New Zealand: (1) the Netherlands Twin Register (NTR)[61,134], (2) E-Risk[64], (3) the Texas Twin Project[66], (4) the AddHealth[62,135], dbGaP accession no. phs001367.v1.p1; (5) WLS[65], dbGaP accession no. phs001157.v1.p1; and (6) the Dunedin Multidisciplinary Health and Development Study[63]. Supplementary Tables 9 and 10 describe

cohort-specific metrics, and we include a short description of the cohorts' populations and recruitment in Supplementary Note. Only participants with European ancestry were included in the analysis, due to the low portability of PGSs between different ancestry populations. PGSs were computed with PLINK based on weights derived using the LD-pred[136] software with an infinitesimal prior and the 1000 Genomes Project phase 3 sample as a reference for the LD structure. LD-pred weights were computed in a shared pipeline to ensure comparability between cohorts. Each outcome (for example, IQ score) was regressed on the Cog and NonCog PGSs and a set of control variables (sex, 10 principal components derived from the genetic data and, for cohorts in which these quantities varied, genotyping chip and age), using Stata 14 for WLS, Stata 15 for E-Risk and the Dunedin Study, and R (versions 3.4.3 and newer) for NTR, AddHealth, and the Texas Twin Project. In cohorts containing related individuals, non-independence of observations from relatives was accounted for using generalized estimation equations (GEE) or by clustering of standard errors at the family level. We used a random effects meta-analysis to aggregate the results across the cohorts. This analysis allows a cohort-specific random intercept. Individual cohort results are in Supplementary Table 11 and meta-analytic estimates in Supplementary Table 12.

Biological annotation

Enrichment of tissue-specific gene expression

We used gene sets defined in Finucane et al.[137] to test for the enrichment of genes specifically expressed in one of 53 GTEx tissues[102], or 152 tissues captured by the Franke et al. aggregation of RNA-sequencing studies[103,104]. This analysis seeks to confirm the role of brain tissues in mediating Cog and NonCog influences on educational attainment. The exact analysis pipeline used is available online (https://github.com/bulik/ldsc/wiki/Cell-type-specific-analyses).

Enrichment of cell-type-specific expression

We leveraged single-cell RNA-sequencing data of cells sampled from the mouse nervous system[107] to identify cell-type-specific RNA expression. Zeisel et al.[107] sequenced cells obtained from 19 regions in the contiguous anatomical regions in the peripheral sensory, enteric, and sympathetic nervous system. After initial quality control, they retained 492,949 cells, which were sampled down to 160,796 high-quality cells. These cells were further grouped into clusters representing 265 broad cell types. We analysed the dataset published by Zeisel et al. containing mean transcript counts for all genes with count >1 for each of the 265 clusters (Supplementary Table 17). We restricted analysis to genes with expression levels above the 25th percentile. For each gene in each cell type, we computed the cell-

type-specific proportion of reads for the gene (normalizing the expression within cell type). We then computed the proportion of proportions over the 265 cell types (computing the specificity of the gene to a specific cell type). We ranked the 12,119 genes retained in terms of specificity to each cell type and then retained the 10% of genes most specific to a cell type as the "cell-type-specific" gene set. We then tested whether any of the 265 cell-type-specific gene sets were enriched in the Cog or NonCog GWAS. This analysis sought to identify specific cell types and specific regions in the brain involved in the aetiology of Cog and NonCog. We further computed the difference in enrichment for Cog and NonCog to test whether any cell types were specific to either trait. For these analyses, we leveraged two widely used enrichment analysis tools: MAGMA[105] and stratified LDSC[106] with the European reference panel from 1000 Genomes Project Phase 3 as SNP location and LD structure reference, Gencode release 19 as gene location reference and the human-mouse homology reference from MGI (http://www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt).

MAGMA

We used MAGMA (v1.07b[105]), a program for gene-set analysis based on GWAS summary statistics. We computed gene-level association statistics using a window of 10 kb around the gene for both Cog and NonCog. We then used MAGMA to run a competitive gene-set analysis, using the gene *p* values and gene-correlation matrix (reflecting LD structure) produced in the gene-level analysis. The competitive gene-set analysis tests whether the genes within the cell-type-specific gene set described above are more strongly associated with Cog/NonCog than other genes.

Stratified LDSC

We used LDSC to compute LD scores for the SNPs in each of our "cell-type-specific" gene sets. Parallel to MAGMA analysis, we added a 10-kb window around each gene. We ran partitioned LDSC to compute the contribution of each gene set to the heritability of Cog and NonCog. To guard against inflation, we used LD score best practices, and included the LD score baseline model (baselineLD.v2.2) in the analysis. We judged the statistical significance of the enrichment based on the *p* value associated with the tau coefficient.

Difference in enrichment between Cog and NonCog

To compute differences in enrichment, we compute a standardized difference between the per-annotation enrichment for Cog and NonCog as:

$$Z_{diff} = \frac{e_{Cog} - e_{NonCog}}{sqrt(se_{Cog}^2 + se_{NonCog}^2 - 2*CTI*se_{Cog}*se_{NonCog})} \quad (1)$$

where e_{Cog} is the enrichment of a particular gene set for Cog, e_{NonCog} is the enrichment for the same gene set for NonCog, se_{Cog} is the standard error of the enrichment for Cog, se_{NonCog} is the standard error of the enrichment for NonCog, and CTI is the LD score cross-trait intercept, a metric of dependence between the GWASs of Cog and NonCog.

We investigated the significance of the difference between Cog and NonCog tau coefficient with Eq. (1) as well as by computing jack-knifed standard errors. From the jack-knifed estimates of the coefficient output by the LDSC software, we computed the jack-knifed estimates and standard errors of the difference between Cog and NonCog tau coefficients, as well as a *z*-statistic for each annotation.

Enrichment of gene expression in the brain

We performed a transcriptome-wide association study (TWAS) using FUSION[108] (http://gusevlab.org/projects/fusion/). We used pre-computed brain-gene-expression weights available on the FUSION website, generated from 452 human individuals as part of the CommonMind Consortium. We then superimposed the bivariate distribution of the results of the TWAS for Cog and NonCog over the bivariate distribution expected given the sample overlap between educational attainment and cognitive performance (the GWAS on which our GWASs of Cog and NonCog are based, see Supplementary Note).

Brain modalities

Brain volumes

We conducted genetic correlation analysis of brain volumes using GWAS results published by Zhao et al.[111], who performed GWASs of total brain volume and 100 regional brain volumes, including 99 grey-matter volumes and total white-matter volume (Supplementary Table 21). Analyses included covariate adjustment for sex, age, their square interaction and 20 principle components. Analyses of regional brain volumes additionally included covariate adjustment for total brain volume. GWAS summary statistics for these 101 brain volumes were obtained from https://med.sites.unc.edu/bigs2/data/gwas-summary-statistics/. Summary statistics were filtered and pre-processed using Genomic-SEM's "munge" function, retaining all HapMap3 SNPs with allele frequency > 0.01 outside the major histocompatibility complex region. We used Genomic-SEM to compute the genetic correlations between Cog, NonCog and brain volumes. Analyses of regional volumes controlled for total brain volume. For each volume, we tested whether correlations differed between Cog and NonCog. Specifically,

we used a chi-squared test to evaluate the null hypothesis that the two genetic correlations were equal. We used FDR adjustment to correct for multiple testing. The FDR adjustment is applied to the results for all grey-matter volumes for Cog and NonCog separately.

White-matter structures

We conducted genetic correlation analysis of white-matter structures using GWAS results published by Zhao et al.[115], who performed GWASs of DTI measures of the integrity of white-matter tracts. DTI parameters were derived for fractional anisotropy, mean diffusivity, axial diffusivity, radial diffusivity, and mode of anisotropy. Each of these parameters was measured for 22 white-matter tracts of interests (Supplementary Table 22), resulting in 110 GWASs. GWAS summary statistics for these 110 GWASs were obtained from https://med.sites.unc.edu/bigs2/data/gwas-summary-statistics/. Summary statistics were filtered and processed using Genomic-SEM's "munge" function, retaining all HapMap3 SNPs with allele frequency > 0.01 outside the major histocompatibility complex region. For each white-matter structure, we tested whether genetic correlations differed between Cog and NonCog. Specifically, we used a chi-squared test to evaluate the null hypothesis that the two genetic correlations were equal. We used FDR adjustment to correct for multiple testing. As these different diffusion parameters are statistically and logically interdependent, having been derived from the same tensor, FDR adjustment was applied to the results for each type of white-matter diffusion parameter separately. FDR correction was applied separately for Cog and NonCog.

ADDITIONAL RESOURCES

A FAQ on why, how and what we studied is available here: https://medium.com/@kph3k/investigating-the-genetic-architecture-of-non-cognitive-skills-using-gwas-by-subtraction-b8743773ce44

A tutorial on how to perform GWAS-by-subtraction:
http://rpubs.com/MichelNivard/565885

Additional resources to Genomic-SEM software:

- A wiki including numerous tutorials:
  https://github.com/MichelNivard/GenomicSEM/wiki
- A Genomic-SEM user group for specific questions relating to models and software:
  https://groups.google.com/g/genomic-sem-users
- A venue to report technical issues:
  https://github.com/MichelNivard/GenomicSEM/issues

CODE AND DATA AVAILABILITY

Code used to run the analyses is available at: https://github.com/PerlineDemange/non-cognitive

A tutorial on how to perform GWAS-by-subtraction: http://rpubs.com/MichelNivard/565885

All additional software used to perform these analyses are available online.

GWAS summary data for NonCog and Cog (excluding 23andMe) have been deposited in the GWAS Catalog with accession nos GCST90011874 and GCST90011875, respectively (NonCog GWAS: ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90011874, Cog GWAS: ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90011875).

For 23andMe dataset access, see https://research.23andme.com/dataset-access/.

Part of the National Longitudinal Study of Adolescent to Adult Health (Add Health) data is publicly available and can be downloaded at the following link: https://data.cpc.unc.edu/projects/2/view#public_li. For restricted access data, details of the data sharing agreement and data access requirements can be found at the following link: https://data.cpc.unc.edu/projects/2/view

The Dunedin study datasets reported in the current article are not publicly available due to lack of informed consent and ethical approval, but are available on request by qualified scientists. Requests require a concept paper describing the purpose of data access, ethical approval at the applicant's university, and provision for secure data access. We offer secure access on the Duke, Otago and King's College campuses. All data analysis scripts and results files are available for review (https://moffittcaspi.trinity.duke.edu/research-topics/dunedin).

The E-Risk Longitudinal Twin Study datasets reported in the current article are not publicly available due to lack of informed consent and ethical approval, but are available on request by qualified scientists. Requests require a concept paper describing the purpose of data access, ethical approval at the applicant's university, and provision for secure data access. We offer secure access on the Duke and King's College campuses. All data analysis scripts and results files are available for review (https://moffittcaspi.trinity.duke.edu/research-topics/erisk).

Netherlands Twin Register data may be accessed, upon approval of the data access committee (email: ntr.datamanagement.fgb@vu.nl).

Researchers will be able to obtain Texas Twins data through managed access. Requests for managed access should be sent to Dr. Elliot Tucker-Drob (tuckerdrob@utexas.edu) and Dr. Paige Harden (harden@utexas.edu), joint principal investigators of the Texas Twin Project.

Wisconsin Longitudinal study data can be requested following this form: https://www.ssc.wisc.edu/wlsresearch/data/Request_Genetic_Data_28_June_2017.pdf

SELECTED SUPPLEMENTARY INFORMATION

Full Supplementary Information and Supplementary Tables can be downloaded at: https://www.nature.com/articles/s41588-020-00754-2#Sec32

Sensitivity test for non-zero correlation of Cog and NonCog

When running GWAS-by-subtraction we assume the model as defined is a correct representation of the relation between cognitive performance (CP) and educational attainment (EA). As the Cog and NonCog latent factors are specified to be uncorrelated, all SNPs that influence Cog will effect both CP and EA, and all SNPs influencing NonCog will only influence EA.

To investigate how a positive non-zero correlation between the Cog and NonCog latent factors could affect results, we re-ran the Genomic-SEM model setting the standardized covariance (i.e. correlation) of NonCog and Cog to 0, 0.1, 0.2, and 0.3. At higher levels of correlation between NonCog and Cog, the NonCog factor explained an increasing percentage of variance in EA: 57% with r_g(Cog, NonCog) = 0 increasing to 78% when r_g(Cog, NonCog) = 0.3. We report the path loadings and the percentage of EA genetic variance explained in Supplementary Table 6.

We next re-estimated genetic correlations of NonCog and Cog with the set of traits in Figure 4 and Supplementary Figure 11. We performed this analysis using models that again set the correlation NonCog and Cog to 0, 0.1, 0.2, and 0,3, as shown in Supplementary Figure 17 for r_g(Cog, NonCog) = 0. Results show a consistent pattern of change in NonCog r_g with target traits. As the correlation of NonCog and Cog is increased, the NonCog r_g with a target trait changes in the direction of the r_g of Cog with that trait. For example, in the case of household income, Cog is positively associated with the trait, therefore as the r_g(Cog, NonCog) increases, the correlation of household income with NonCog increases positively (at r_g(Cog, NonCog) = 0 the r_g is 0.61, at r_g(Cog, NonCog) = 0.3 it is 0.76). These changes are small in magnitude, but sometimes alter the statistical significance of the r_g. When r_g(Cog, NonCog) is set to 0.3, the following r_gs with NonCog, which were not statistically different from zero under the r_g(Cog, NonCog) = 0 specification, become statistically significant: Age at menopause, Autism Spectrum Disorder and Chronotype. In contrast, only Conscientiousness and Self-report empathy had a statistically significant r_g with NonCog under r_g(Cog, NonCog) = 0 but were not significantly genetically correlated with NonCog when when r_g(Cog, NonCog) = 0.3. Results are presented in Supplementary Table 7 and Supplementary Figure 5.

The magnitude of the change of the genetic correlation with NonCog is dependent on the genetic correlation of Cog with the trait: the stronger the genetic correlation with Cog, the bigger increase/decrease of the genetic correlation with NonCog (Supplementary Figure 6).

Sensitivity test for causal relation between CP and EA

Our primary model (Figure 1) assumes all genetic effects on CP also affect EA. This assumption is reasonable here, as cognitive ability is an important driver of educational success. In fact, many high-income countries mandate cognitive test-scores as entry to higher education. Furthermore, tests of polygenicity consistently find that a smaller portion of the genome has an effect on CP then on EA, consistent with a model where CP causes EA.

However, we can consider a violation of our assumed model based on reasonable estimates from the literature. Ritchie and Tucker-Drob (2018)[24] find across multiple studies, which rely on control variables or natural experiments, that there is a robust but small effect of education on IQ. Consistent with these results, Savage et al. (2018)[138] performed a GWAS of intelligence and, using Mendelian randomization, found a bidirectional effect between IQ and educational attainment.

We investigated the impact of a reciprocal effect of EA on CP on our results. We can only allow for, but not estimate, such an effect in the context of our model, as the effect is not identified (Supplementary Figure 7); we chose a small standardized effect size of 0.2. Based on this alternative model, we reanalysed the genome-wide significant SNPs for Cog and NonCog and found minimal change in Z-statistics (see Supplementary Figure 8). We further re-computed the genetic correlations between Cog and NonCog and the external traits in Figure 4 and Supplementary Figure 11 (by adapting the model Supplementary Figure 17). We observe minimal changes in the genetic correlations as well (Supplementary Figure 9 and Table 8). Therefore, our results appear robust to the relaxation of the assumption that the primary causal relationship is from CP to EA and not vice versa.

CHAPTER 3

ESTIMATING EFFECTS OF PARENTS' COGNITIVE AND NONCOGNITIVE SKILLS ON OFFSPRING EDUCATION USING POLYGENIC SCORES

Published as: **Demange, P.A.**, Hottenga, J.J., Abdellaoui, A., Eilertsen, E.M., Malanchini, M., Domingue, B.W., Armstrong-Carter, E., de Zeeuw, E.L., Rimfeld, K., Boomsma D.I., van Bergen, E., Breen, G., Nivard, M.G., & Cheesman, R. (2022). Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. Nature Communications 13, 4801 https://doi.org/10.1038/s41467-022-32003-x

Supplementary materials accessible at:

https://www.nature.com/articles/s41467-022-32003-x#Sec32

ABSTRACT

Understanding how parents' cognitive and noncognitive skills influence offspring education is essential for educational, family and economic policy. We use genetics (GWAS-by-subtraction) to assess a latent, broad noncognitive skills dimension. To index parental effects controlling for genetic transmission, we estimate indirect parental genetic effects of polygenic scores on childhood and adulthood educational outcomes, using siblings (N = 47,459), adoptees (N = 6,407), and parent-offspring trios (N = 2,534) in three UK and Dutch cohorts. We find that parental cognitive and noncognitive skills affect offspring education through their environment: on average across cohorts and designs, indirect genetic effects explain 36-40% of population polygenic score associations. However, indirect genetic effects are lower for achievement in the Dutch cohort, and for the adoption design. We identify potential causes of higher sibling- and trio-based estimates: prenatal indirect genetic effects, population stratification, and assortative mating. Our phenotype-agnostic, genetically sensitive approach has established overall environmental effects of parents' skills, facilitating future mechanistic work.

INTRODUCTION

Parents and children tend to have similar educational outcomes[1]. Since education is highly predictive of social mobility and health across the lifespan[139,140], understanding the mechanisms underlying the intergenerational transmission of education could inform efforts to alleviate inequalities. Many studies have investigated how much certain parental characteristics influence offspring education, but relatively few have considered noncognitive skills. The term 'noncognitive' describes skills that differ from what has traditionally been education's primary focus: academic and cognitive performance. The umbrella of noncognitive skills encompasses a wide range of competencies, including academic motivation, perseverance, mindsets, learning strategies, and social skills[29,141]. Cognitive skills like executive functioning, working memory, and verbal IQ are more precisely integral to cognitive functioning, but both cognitive and noncognitive skills are critical for educational success[29]. Research in developmental psychology[142], economics[143], and sociology[144] has suggested that parents socially influence their children's noncognitive skills including emotion regulation, social capacities, attitudes and motivations[33,34]. Given that noncognitive skills (particularly self-control and emotion regulation[26,145]) support education, it follows that parents' noncognitive skills may also affect children's educational outcomes.

Prior research has detected small associations between measured parental noncognitive skills and offspring educational outcomes. In one study, mothers' locus of control was the only significant noncognitive predictor of offspring college attendance (β = 0.02, p<0.05; β = ~0.01 for maternal self-concept and self-esteem, both non-significant)[146]. Mothers' cognitive skills, measured by the U.S. Armed Forces Qualifying Test, were a stronger predictor (β = 0.06, p<0.01). Another study found that fathers' noncognitive skills were associated with sons' standardised test scores at age 16 (β = 0.09)[147]. Here, noncognitive skills were measured by a single composite of extraversion, neuroticism, persistence, and perseverance from a standardised Swedish military-oriented psychological evaluation. Additionally, parents' attitudes towards education and social skills have been found to account for 8% of the socioeconomic gap in children's achievement[148]. The contributions of specific measured parental traits that were included were also not stated.

Two key limitations weaken this base of evidence on the effects of parents' skills on offspring education: challenges with phenotypic assessments of parents' noncognitive skills, and genetic confounding.

First, regarding assessment, whereas cognitive skills can be directly measured by tests of domain-specific or general cognitive performance, noncognitive skills are more challenging to capture[30,149]. There is little agreement on what noncognitive skills to measure. Some

researchers focus on personality, whereas others include self-control, self-esteem, motivation, and interests. Importantly, studies identifying partial effects of specific parental cognitive and noncognitive skills are less informative about the overall influences of these domains. Measurement error could also mean that effects of parents' noncognitive skills have been underestimated.

Genetic methods offer an alternative approach to defining parents' noncognitive skills. Both cognitive and typically-studied noncognitive skills are substantially genetically influenced, with twin study heritability estimates of 40-70%[150,151]. A new method – 'GWAS-by-subtraction' – makes it possible to assess a broad latent genetic noncognitive construct, by 'subtracting' cognitive ability-related genetic variation from educational attainment genetic variation[152]. This follows an influential definition of noncognitive skills from economics[32] as all traits positively contributing to educational and professional success that are not cognitive skills. This noncognitive genetic construct — which could otherwise be conceptualized as 'not-cognitive' — is associated with higher socioeconomic attainment, more open and conscientious personality, and some psychiatric disorders (e.g. higher risk for schizophrenia, lower risk for attention deficit/hyperactivity disorder). In the present study, we use this GWAS-by-subtraction measure of noncognitive skills to capture the overall effect of all noncognitive parent phenotypes on offspring education. This phenotype-agnostic approach is somewhat loose: it could include parental phenotypes not traditionally classed as 'noncognitive' or 'skills'. However, it provides a useful first step towards characterizing pathways from specific parental skills to offspring educational outcomes. After establishing overall effects, complementary research designs using measured parental noncognitive skills can subsequently be used to identify specific mediating mechanisms.

A second challenge is to distinguish social (i.e. environmental) from genetic transmission. None of the associations between parental skills and offspring education cited above were estimated using genetically sensitive designs. This is problematic, because from just parent-offspring correlations one cannot conclude that parents' skills shape offspring education, for instance by providing resources, experiences, and support. Ignoring any shared genetic influences on parents' skills and child educational outcomes confounds estimation of the effects of parental phenotypes on offspring outcome[35]. To establish the extent that parents' (non-)cognitive skills influence child educational outcomes socially, it is vital to control for inherited genetic effects.

Genetic study designs can isolate environmental effects of parental skills on offspring education, controlling for genetic transmission. Several designs estimate a genetic effect of the child's genotype on the child phenotype (direct genetic effect), and an environmentally

mediated effect of the parental genotype on the child's phenotype (parental indirect genetic effect). For example, non-transmitted genetic variants affect offspring phenotypes indirectly via the environment shaped by parental phenotypes[153,154]. Polygenic scores (individual-level indices of trait-specific genetic endowment; PGS) for educational attainment based on parents' non-transmitted variants, are associated with offspring attainment[155–157]. Complementary evidence of indirect effects of parents' education-linked genetics on offspring education has also accumulated from sibling and adoption PGS designs[155,156,158,159]. To obtain estimates of indirect genetic effects using sibling data, within-sibling genetic associations (first developed to estimate direct genetic effects independent of population biases[160,161]) are compared to population-based associations. To obtain estimates of indirect genetic effects using adoption data, genetic associations estimated for adoptees and non-adopted individuals are compared[159]. Notably, variance decomposition as well as PGS methods can be applied to disentangle direct and indirect genetic effects, but the former requires much larger sample sizes[162–165]. It is not known whether parental indirect genetic effects on offspring education occur through cognitive or noncognitive pathways (or both), because studies have not parsed out the contributions of sub-components of the educational attainment PGS.

Here, we directly compare estimates of parental indirect genetic effects obtained from different designs. Estimation of genetic associations may involve numerous biases[166–168]. Sibling, adoption, and non-transmitted allele designs have different assumptions and subtle differences in biases and components affecting the estimated indirect genetic effect. As shown by our data simulations indirect genetic effect estimates from the adoption design may be less biased by population stratification and assortative mating than the sibling and non-transmitted allele designs (see Supplementary Note 6 and our GitHub repository[169]). However, estimates obtained from the adoption design do not capture prenatal parental environmental effects on child education and may be less generalisable to the population. The sibling design may estimate parental indirect genetic effects with more bias from sibling genetic effects. Triangulation across designs and sensitivity analyses can help detect possible biases and quantify parental indirect genetic effects and other environmental effects[13,167].

In the current study (pre-registration: https://osf.io/mk938/), we use a novel approach to estimate the social effects of parents' cognitive and noncognitive skills on offspring education. We deploy GWAS-by-subtraction to estimate individuals' genetic endowments (PGS) for cognitive and noncognitive skills, and test how much these operate environmentally via parental influences on offspring educational outcomes. We provide a comparison of

parental indirect genetic effects in three cohorts of genotyped families in two countries (UK Biobank, UK Twins Early Development Study, Netherlands Twin Register). Each cohort includes multiple achievement outcome measures (i.e. standardised test results and teacher-reported grades in childhood and adolescence) and attainment (i.e. years of completed education reported in adulthood). We triangulate across three complementary study designs for estimating parental indirect genetic effects and assess the presence of components and biases.

RESULTS

GWAS-by-subtraction results

We identified the genetic components of cognitive and noncognitive skills using Genomic-SEM, following Demange et al.[152], in samples that excluded participants used for polygenic score analyses. Educational attainment and cognitive performance meta-analytic summary statistics (see Methods) were regressed on two independent latent variables, Cog and NonCog (see Supplementary Figure 1). These two latent factors were then regressed on 1,071,804 HapMap3 SNPs in a genome wide association (GWA) design. The LD score regression-based SNP heritabilities of Cog and NonCog were 0.184 (*SE* = 0.007) and 0.054 (*SE* = 0.002), respectively. More information on the GWASs is presented in Supplementary Data 1.

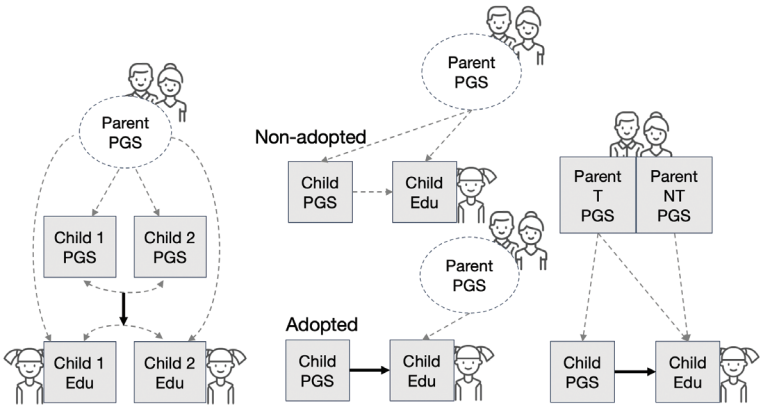
Descriptive statistics

SNP associations with the Cog and NonCog latent variables provided the weights to create individual-level polygenic scores in 3 cohorts with family data and educational achievement and/or attainment outcomes. Sample sizes for individuals with polygenic score and educational outcome data were: 39,500 UK Biobank siblings, 6,409 UK Biobank adoptees, up to 4,796 DZ twins in the Twins Early Development Study (TEDS), up to 3,163 twins and siblings in the Netherlands Twin Register (NTR), and up to 2,534 NTR individuals with both parents genotyped. Full phenotypic descriptive statistics are available in Supplementary Data 2.

Overview of three family-based polygenic score designs

To estimate direct offspring-led and indirect parent-led effects of PGS for cognitive and noncognitive skills on educational outcomes, we considered three family-based genomic designs. The designs are illustrated in Figure 1. All models jointly included Cog and NonCog PGS. Note that population effects are equivalent to PGS effects estimated in standard

population analyses that do not use within-family data. In contrast, within-family designs exploit the principles of Mendelian segregation or the natural experiment of adoption to separate direct and indirect/social components of the overall population PGS effect. Importantly, a direct genetic effect is only direct in the sense that it does not originate from another individual's genotype. Direct effects are also not 'purely' genetic, but lead to educational outcomes via intermediate pathways, and are expressed in the context of environments.



	Sibling design	Adoption design	Non-transmitted PGS design
Population genetic effect	$\beta_{Population}$	$\beta_{Non\text{-}adopted}$	$\beta_{Transmitted}$
Direct genetic effect	$\beta_{Within\text{-}siblings}$	$\beta_{Adopted}$	$\beta_{Transmitted} - \beta_{Non\text{-}transmitted}$
Indirect genetic effect	$\beta_{Population} - \beta_{Within\text{-}siblings}$	$\beta_{Non\text{-}adopted} - \beta_{Adopted}$	$\beta_{Non\text{-}transmitted}$

Figure 1 | Analytical designs to estimate direct and parental indirect genetic effects. Note: square = observed variable, circle = unobserved / latent variable; β = estimated effect of polygenic score (PGS) on outcome; the population effect of a PGS captures both direct and indirect genetic effects; direct genetic effects (controlling for indirect genetic effects) are represented with solid arrows. Icons made by Freepik from www.flaticon.com.

First, the sibling design estimates indirect genetic effects by comparing population-level and within-family (i.e. within-sibling or within-DZ twin) PGS associations (Eq. (1))[158]. The direct effect of a polygenic score is estimated based on genetic differences between siblings, which are due to random segregations of parental genetic material, independent of shared family effects (including parental indirect genetic effects). Specifically, the direct effect is estimated using a variable representing individuals' (i) polygenic scores minus the average polygenic score for their family (j): the within-family beta (β_{Within} in Eq. (1)). The population effect of a polygenic score is estimated in a separate model, simply regressing the outcome variable on polygenic score differences between individuals from different families (Eq. (2)).

The indirect genetic effect is obtained by subtracting the within-family PGS effect estimate from the population effect estimate.

$$\begin{aligned} EA_{ij} = \alpha_0 0 &+ \beta_{Within_{Cog}}\left(PGS_{Cog_{ij}} - \overline{PGS_{Cog_j}}\right) + \beta_{Between_{Cog}}\left(\overline{PGS_{Cog_j}}\right) \\ &+ \beta_{Within_{NonCog}}\left(PGS_{NonCog_{ij}} - \overline{PGS_{NonCog_j}}\right) \\ &+ \beta_{Between_{NonCog}}\left(\overline{PGS_{NonCog_j}}\right) + Z_{ij} \end{aligned} \quad (1)$$

$$EA_{ij} = \alpha_0 0 + \beta_{Cog}\left(PGS_{Cog_{ij}}\right) + \beta_{NonCog}\left(PGS_{NonCog_{ij}}\right) + Z_{ij} \quad (2)$$

Note: EA is the educational outcome, $\overline{PGS}$ is the polygenic score (for Cog PGS_{Cog} and NonCog PGS_{NonCog}). PGS refers to the average polygenic score in the family j. i refers to the individual sibling. α0 refers to the intercept, Z are covariates for the individual i: sex, age, sex*age, the first 10 principal components, and genotyping platform. See Supplementary Note 5 for a comparison of different versions of this sibling design, using data simulations.

Second, indirect genetic effects can be estimated by comparing polygenic score associations estimated in a sample of adoptees against those estimated for individuals who were reared by their biological parents[159]. Therefore, we estimate the regression model shown in Eq. (2) separately for adoptees and for non-adopted individuals.

The population effect is estimated as the polygenic score effect on phenotypic variation among non-adopted individuals (i.e. a combination of direct and indirect genetic mechanisms). The direct genetic effect is the effect of the polygenic score among adoptees. Adoptees do not share genes by descent with their adoptive parents, so we expect their polygenic scores to be uncorrelated with the genotypes of their adoptive parents. Therefore, the polygenic score effect in adoptees cannot be inflated by environmentally mediated parental indirect genetic effects. In this design, the indirect genetic effect is estimated by subtracting this direct PGS effect from the population effect estimated in the non-adopted group. When taking the difference, it is important that the groups are similar in terms of all observed and unobserved confounders, an untestable assumption that is unlikely to always hold. We found small differences between adoptees and non-adopted individuals in the UK Biobank in their demographic and early-life characteristics. Cohen's d values were $d<0.15$ for Cog and NonCog PGS and educational attainment, and $d = 0.31$ for the birth weight. The pattern of geographical clustering of adopted and non-adopted participants across the UK was highly similar (see Supplementary Data 11, Supplementary Note 3, and Supplementary Figure 2).

Third, indirect genetic effects can be estimated, and disentangled from direct genetic effects, using information on parental genetic variation that was not transmitted to offspring[155,156] (Eq. (3)).

$$\begin{aligned} EA = {} & \alpha_0 0 + \beta_{T_{Cog}}\left(PGS_{T_{Cog}}\right) + \beta_{T_{NonCog}}\left(PGS_{T_{NonCog}}\right) \\ & + \beta_{NT_{Cog}}\left(PGS_{NT_{Cog}}\right) + \beta_{NT_{NonCog}}\left(PGS_{NT_{NonCog}}\right) + Z \end{aligned} \quad (3)$$

The population effect is estimated from a polygenic score based on transmitted variants (β_T). Transmitted genetic variants are present in an offspring and in at least one of their parents, and so may influence offspring education via both direct and indirect mechanisms. The parental indirect genetic effect is estimated as the effect of a polygenic score based on parental variants that were not transmitted to offspring (β_{NT}). Non-transmitted variants can only take effect on offspring education through the environment. The direct genetic effect is estimated by partialling out the effect of the non-transmitted polygenic score from that of the transmitted polygenic score (β_T - β_{NT}). Maternal and paternal scores are averaged to create overall parental non-transmitted polygenic scores. We did not distinguish between maternal and paternal PGS, due to the replicated evidence that mothers' and fathers' PGS for educational attainment have equal effects on offspring education[170,171], and to enable closer comparison with the adoption and sibling designs, which yield estimates of the overall parental genetic effect. Notably, regressing offspring phenotype on offspring PGS and parental PGS would allow equivalent estimation of the parental indirect genetic effect without haplotype estimation[172].

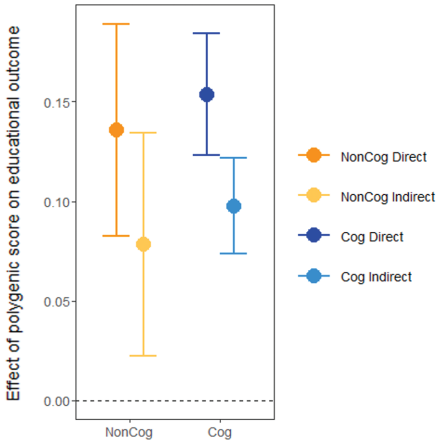
Parents' heritable cognitive and noncognitive skills environmentally influence offspring education.

We investigated environmental effects of parents' noncognitive and cognitive skills on offspring education by estimating parental indirect genetic effects of NonCog and Cog PGS. Figure 2a shows that, for both NonCog and Cog PGS, indirect genetic effects of parents on offspring education were present (meta-analytic indirect β_{NonCog} = 0.08, *SE* = 0.03; indirect β_{Cog} = 0.10, *SE* = 0.01), in addition to direct genetic effects (direct β_{NonCog} = 0.14, *SE* = 0.03; direct β_{Cog} = 0.15, *SE* = 0.02). Averaged across all designs, outcomes and cohorts, indirect environmentally mediated effects explained 36% of the population effect of the NonCog PGS, and 40% of the population effect of the Cog PGS. However, results varied depending on the methods used and outcomes investigated. Results per cohort, outcome and design, as well as population genetic effects and the ratio of indirect to population effects are reported in Supplementary Data 3 and Supplementary Figure 3, 4 and 5. Meta-analytic results are reported in Supplementary Data 4. Z-tests results comparing direct and indirect effects are reported Supplementary Data 5.

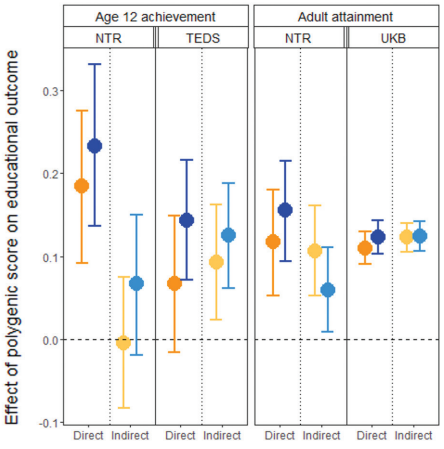
Estimates of indirect genetic effects vary by age, outcome and cohort

Figure 2b shows estimates of direct and indirect genetic effects of NonCog and Cog PGS for different cohorts and educational outcomes, holding the design constant (i.e. the sibling design, which was available for all cohorts and outcomes). Estimates were highly consistent across cohorts except for age 12 achievement in Dutch versus UK cohorts: indirect genetic effects were negligible and represented a small fraction of the population effect in NTR (3% and 23% for NonCog and Cog, respectively), whereas they accounted for 56% and 48% of the population effects of NonCog and Cog PGS in TEDS. For adult educational attainment, estimates of direct and indirect effects were more similar for the Dutch (NTR: indirect β_{NonCog} = 0.11, *SE* = 0.03; indirect β_{Cog} = 0.06, *SE* = 0.03) and UK (UKB: indirect β_{NonCog} = 0.12, *SE* = 0.01; indirect β_{Cog} = 0.12, *SE* = 0.01) cohorts. See Supplementary Data 3 for full results.

a. Meta-analytic results



b. Sibling design by cohort



c. Educational attainment by design

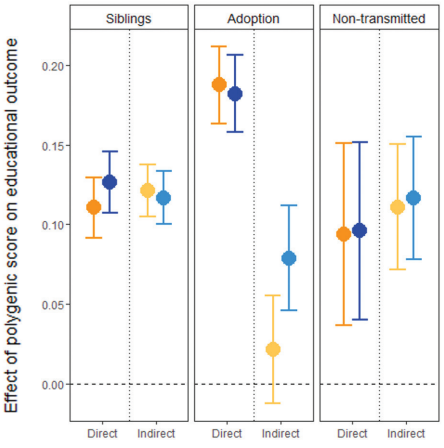


Figure 2 | Estimated direct and indirect genetic effects of NonCog and Cog PGS on educational outcomes. a. Meta-analytic results. Meta-analysed estimates of direct and indirect genetic effects of NonCog and Cog PGS on educational outcomes (N = 68,919). Indirect genetic effects work through the environment that parents provide for their children. Notes: beta coefficients were obtained from meta-analysis of effects across cohorts, designs and outcome phenotypes; bars = 95% CIs. **b. Sibling design by cohort.** Estimates of direct and indirect effects of NonCog and Cog PGS by cohort (for age 12 and adult outcomes), using the sibling design only. NTR is a Dutch cohort (N = 1631 and N = 3163 respectively), TEDS (N = 2862) and UKB (N = 16,624) are UK cohorts; bars = 95% CIs. **c. Educational attainment by design.** Estimates of direct and indirect effect of NonCog and Cog PGS by analytic design (for adult educational attainment outcomes only). Samples sizes: N = 42,663 (results meta-analysed across UKB and NTR); N = 6407 adoptees and 6500 non-adopted individuals (UKB); N = 2534 trios in NTR; bars = 95% CIs.

Estimates of indirect genetic effects depend on the analytical design

Figure 2c shows estimates of direct and indirect genetic effects of NonCog and Cog PGS for different designs, holding the phenotype constant (i.e. educational attainment, which was available for all three methods). While estimates obtained with sibling and non-transmitted PGS methods indicate equal indirect effect sizes (indirect βs for educational attainment ranged between 0.11-0.12; see Supplementary Data 3 and 4), the adoption design yielded low to null indirect genetic effects for both NonCog and Cog PGS (indirect β_{NonCog} = 0.02, *SE* = 0.02; indirect β_{Cog} = 0.08, *SE* = 0.02).

Figure 3 summarises how the three designs estimate parental indirect genetic effects in the presence of different contributors, thus highlighting possible explanations for lower adoption-based estimates. This information is based on simulations (see Supplementary Notes 4 and 6, Supplementary Figure 9, and our GitHub repository[169]). We consider prenatal and postnatal parental indirect genetic effects as components of the total parental indirect genetic effect, and other simulated contributors as biases. First, unlike the sibling and non-transmitted allele designs, the adoption design does not capture indirect genetic effects occurring in the prenatal period. Second, the adoption design estimates indirect genetic effects with less bias from population stratification. Third, the adoption design estimates indirect genetic effects with less bias from assortative mating than the sibling design, and, most likely, than the non-transmitted alleles design. How the bias in the adoption design estimates compares to the non-transmitted design depends on the precision of the polygenic score, see Supplementary Information. Any excess indirect genetic effect estimated in the sibling/non-transmitted allele designs compared to the adoption design therefore indicates the overall impact of prenatal indirect genetic effects, population stratification, and assortative mating.

With the adoption design, the indirect genetic effect of the NonCog PGS on educational attainment in UK Biobank is 83% lower than with the sibling design, while it is only 33%

lower for Cog. This suggests that estimates for NonCog are affected more strongly than Cog by population stratification, assortative mating and/or prenatal indirect genetic effects.

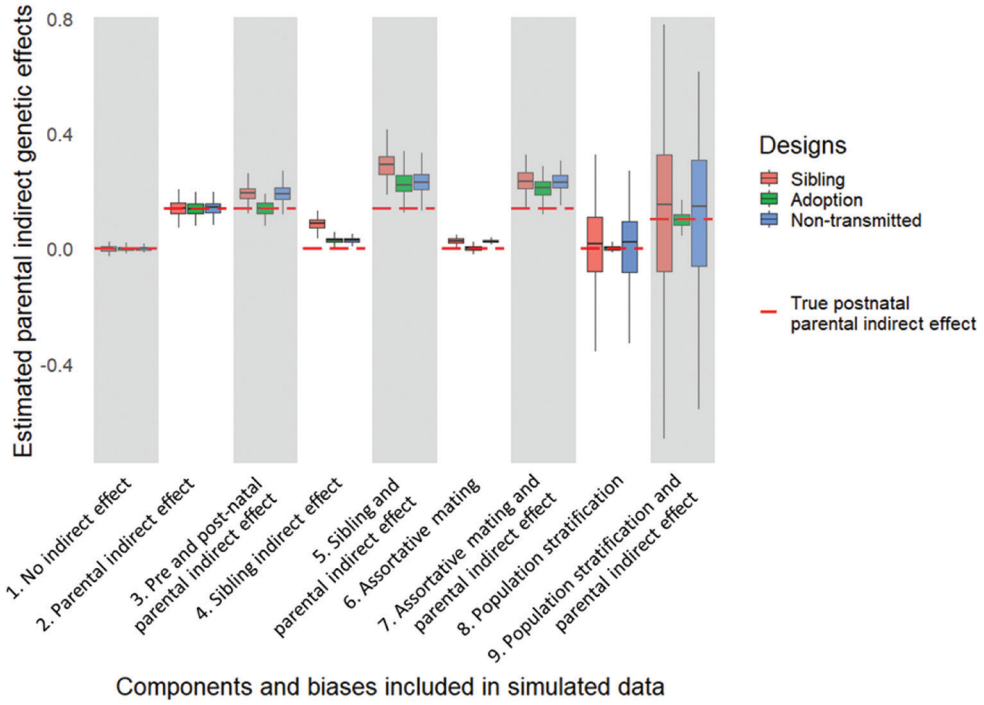


Figure 3 | Estimates of parental indirect genetic effects from the three designs, based on data simulated to include different components and biases. Components include parental prenatal and postnatal indirect genetic effects. Biases include sibling indirect genetic effects, assortative mating, and population stratification. Boxplots of 100 replicates based on a simulated sample of 20,000 families. The centre line represents the median, the box limits are the 1st and 3rd quartile, and the whiskers reach the maximum value within 1.5 times the interquartile range. Outlying values are not represented. For clarity, the red line benchmarks the true simulated postnatal parental indirect effect, although we note that *prenatal* parental genetic effects are a component rather than a bias of the parental indirect genetic effect.

Population phenomena may inflate indirect genetic effect estimates

Although triangulating designs suggested that prenatal indirect genetic effects, population stratification, and assortative mating may contribute to the higher estimated parental indirect genetic effects from non-transmitted alleles/sibling designs relative to the adoption design, this approach cannot disentangle the relative importance of these individual biases. To this end, we conducted additional sensitivity analyses to assess the magnitudes of these biases (not pre-registered).

First, we analysed the GWAS summary data on which the polygenic scores were based,

using LD score regression to detect population stratification. The LD score regression ratio statistics of uncorrected educational attainment and cognitive performance GWASs were 0.11 (*SE* = 0.01) and 0.06 (*SE* = 0.01), respectively (Supplementary Data 1). These non-null estimates indicated that a small but significant portion of the GWAS signal was potentially attributable to residual population stratification. As CP seems less prone to population stratification than EA, it is possible our estimates of direct and indirect genetic effects of NonCog were more biased by population stratification than Cog.

Second, we detected slight evidence of assortative mating, which appeared stronger in the UK than Dutch cohorts. In NTR, parental PGS correlations are non-significant (NonCog r = 0.03, Cog r = 0.02). Sibling PGS intraclass correlations ranged between 0.49-0.52 in NTR, and between 0.53-0.56 in TEDS and UK Biobank. This supports the presence of assortative mating on NonCog and Cog PGS potentially biasing our estimates of indirect genetic effects in UK cohorts, but less in our Dutch cohort. See Supplementary Data 6 for full correlations.

Third, our sensitivity analyses did not support the presence of indirect effects of siblings' NonCog and Cog PGS on individuals' educational outcomes. Our first approach leveraged sibling polygenic scores, the rationale being that in the presence of a sibling effect, a sibling's PGS will influence a child's outcome beyond child and parent PGS. In NTR, siblings' NonCog or Cog PGS had non-significant effects (Supplementary Data 7). In a second approach, the difference in PGS effects on EA between monozygotic (MZ) and dizygotic (DZ) individuals was tested. Since MZ twins are more genetically similar than DZ twins, their PGS should capture more of the indirect genetic effect of their twin. In NTR and TEDS, PGS effects were not significantly different between MZs and DZs (Supplementary Data 8 & Supplementary Figure 6). Finally, in UKB, we tested PGS effects on EA given the number of siblings individuals reported having. If more siblings lead to a stronger sibling effect, this will be captured as an increased effect of an individual's own PGS on the outcome in the presence of more genetically related siblings. As a negative control, we conducted the same analysis in adoptees. Since adoptees are unrelated to their siblings, their PGS do not capture sibling effects at any family size. NonCog PGS effects weakly increased with number of siblings, but this pattern was also present in adoptees, suggesting confounding by unobserved characteristics of families with numerous children (Supplementary Data 9 & Supplementary Figure 7).

DISCUSSION

We used genetic methods to study environmental effects of parents' skills on child education. We found evidence that characteristics tagged by NonCog and Cog polygenic scores (PGS)

are both involved in how parents provide environments conducive to offspring education. Indeed, indirect genetic mechanisms explained 36% of the population effect of the NonCog PGS, and 40% of the population effect of the Cog PGS (population β_{NonCog} = 0.22, *SE* = 0.01; population β_{Cog} = 0.25, *SE* = 0.01). This result was consistent across countries, generations, outcomes, and analytic designs, with two notable exceptions. First, estimated parental indirect genetic effects were null for childhood achievement in our Dutch cohort (NTR), but not for comparable outcomes in our UK cohort (TEDS). Second, parental indirect genetic effects estimated with the adoption design were lower than for the sibling and non-transmitted allele designs, particularly for the NonCog PGS. Given our evidence from data simulations that the adoption-based estimates of indirect genetic effects do not account for prenatal effects and may be more robust to population stratification and assortative mating, these factors may contribute substantially to estimates from the other two designs, especially for the NonCog PGS. This was supported by results from sensitivity analyses.

This study demonstrates utility of genetic methods for assessing elusive phenomena: noncognitive skills, and genuine environmental influences from parents, unconfounded by offspring-led effects of inherited genes. Compared to analysing a set of measured parental noncognitive skills, our GWAS-by-subtraction approach captures a wider array of traits linked genetically to attainment, and therefore broadly quantifies the overall salience of parents' noncognitive skills. Our evidence that parents' noncognitive and cognitive skills are both important for children's education complements the growing literature that has considered effects of specific measured skills within both of these domains[146,147]. These studies found that effects of parents' noncognitive skills on offspring education were less than half the size of the effects of parents' cognitive skills. In contrast, we found that indirect genetic effects of NonCog PGS were almost as large as for Cog skills. This discrepancy might stem in part from our comprehensive definition of noncognitive skills, as we do not rely on possibly unreliable and incomplete phenotypic measures. Importantly, the parental indirect genetic effects we have identified may capture proximal forms of 'nurture' (e.g. a parent directly training their child's cognitive skills, or cultivating their child's learning habits through participation and support) and/or more distal environmental effects (e.g. a parent's openness to experience leading them to move to an area with good schools). The environmental effects of parents' noncognitive and cognitive skills are likely to be larger than we estimate, because our approach only captures effects of parent skills tagged by current GWAS. Polygenic scores index a subset of the common genetic component of parent skills, which is in turn a fraction of the total genetic component (missing heritability[173,174]), and cannot account for the non-heritable component of parent skills.

The lower importance of parental indirect genetic effects for child achievement in the Netherlands compared to the UK indicates that our UK achievement outcomes more strongly capture variation in family background. This difference could result from the design of these achievement measures: Dutch test results are standardized based on a representative population, but UK teacher reports might still be affected by student social background. Societal differences between the two countries might offer another explanation, as indirect genetic effects might be seen as indicator of social inequality (similarly to shared-environment variance in twin studies[175]). For adult attainment, results were more similar across UK and Dutch cohorts, corresponding with recent evidence for consistent shared-environment influence on educational attainment across social models[176]. This consistency also suggests that the difference in childhood is not due to a cohort or population difference. The higher indirect genetic effects in adult attainment in the Netherlands might reflect an increase in environmental variance following tracking taking place in secondary schools[157]. Indeed, socioeconomic disparities in achievement seem to increase more between ages 10 and 15 in the Netherlands than in the UK[177] and children whose parents have a higher education are more likely to enrol in a higher educational track independently of their achievement at age 12[178], suggestive of greater parental effects on secondary and later education, which should be tested in further studies.

We found that the design used to estimate indirect genetic effects matters, with the adoption design giving systematically lower estimates. Direct comparison of results across designs suggested that 33% (for Cog) and 83% (for NonCog) of the indirect genetic effects on adult educational attainment, estimated using the sibling design, are at least in part due to population stratification, assortative mating, and prenatal indirect genetic effects. The importance of population stratification for genetic associations with educational attainment was suggested by recent UK Biobank studies[179,180]. Our sensitivity analyses also indicated residual population stratification, which was more severe for the NonCog GWAS. There was some evidence of assortative mating, with sibling PGS correlations above expectation (>0.5) particularly in the UK cohorts. This country difference in assortment is supported by the lower estimated spouse PGS correlations in NTR (0.02 for Cog, 0.03 for NonCog) than for the EA PGS in the UK Biobank (0.06)[181]. There was no statistically significant difference in assortative mating between Cog and NonCog, suggesting that population stratification explains the particularly large design-based discrepancy between estimates of indirect genetic effects for NonCog (but possibly also differential bias in the Cog versus NonCog GWASs; see Limitations). Population stratification should be carefully considered in studies using NonCog PGS. Structural equation models, leveraging within-family polygenic scores and phenotypes, are being developed to parse the contributions of indirect and direct genetic

effects to complex traits from assortative mating (both disequilibrium and equilibrium forms) and population stratification[182,183]. Another consideration for future research is that indirect genetic effects on education might span across more than a single generation, for example the influence of grandparents. Since cumulative indirect genetic effects are all removed when a child is adopted, their presence would contribute to the observed difference in indirect effect between the adoption and other designs.

Regarding siblings, we did not find evidence that indirect effects of siblings' NonCog and Cog PGS affect individual differences in educational outcomes, using three different approaches. This corresponds with null findings regarding indirect effects of siblings' educational attainment genetics in the UK Biobank[179,180]. However, other UK Biobank studies have detected indirect effects of older siblings' EA PGS on younger siblings' educational attainment[184], and parental compensation for sibling EA PGS differences[185], suggesting that more subtle analyses are required to understand sibling effects. There is also some evidence for sibling effects on educational attainment in other populations, based on the EA PGS[156] and on extended twin family data[186]. It is possible that our PGS analyses were not sufficiently powered to detect indirect genetic effects of siblings, since they were based on lower sample size than our main analyses. However, our results suggest that indirect genetic effects of siblings on education are small. Therefore, our methods provide good proxies for parental indirect genetic effects, with minimal inflation from sibling effects.

Our data suggest that the adoption design may provide a useful lower-bound estimate of indirect genetic effects of parents. Given that there was no evidence for sibling effects of the Cog or NonCog PGS, our adoption-based estimates, which appear to be less biased by population stratification and assortative mating, should give a closer measure of (postnatal) parental indirect genetic effects in the absence of other issues. However, adoptees and non-adopted individuals differ in unobserved and observed ways, including birthweight (d = 0.3). These differences likely make adoption-based estimates of indirect genetic effects, which rely on a comparison of the two groups, less reliable. Moreover, three additional factors may make the adoption-based estimates of indirect genetic effects too conservative. First, adoption based indirect effect estimates exclude prenatal indirect genetic effects (and indirect genetic effects taking place between the birth and moment of adoption), which might influence educational outcomes[187,188]. While we are unable to test for prenatal indirect effects, these could be investigated in cohorts with pregnancy information, adjusting for postnatal indirect genetic effects. Second, adoptees may have been exposed to a narrower range of environments (e.g. family socioeconomic status) compared to non-adopted individuals[189]. This form of selection bias is likely to increase the genetic variance at the expense of the

indirect genetic effect. Third, selective placement of children in adoptive families matching characteristics of their biological families, or adoption of children by close relatives[190], could result in correlation between child and (adoptive) parent genotypes, leading to an underestimation of the indirect genetic effect. There is modest evidence for selective placement of adoptees based on education in the US[191]. We cannot control for selection and relatedness (e.g. by excluding individuals who were adopted by relatives and/or adopted relatively late in development), since there is no information on the adoptive parents in the UK Biobank resource.

We acknowledge several limitations. First, while we suggest that an attribute of our study is the broad and phenotype-agnostic characterisation of noncognitive skills, our GWAS-by-subtraction approach is unable to identify specific parental characteristics and is also still limited by measures of cognitive performance and educational attainment in the original GWASs. Some cognitive skills might not be reflected in the available Cognitive Performance GWAS, so the NonCog factor could capture genetic influences affecting cognition. However, previous analyses have shown that a NonCog PGS based on GWAS-by-subtraction predicts substantially less variation in cognition than the Cog PGS[152]. Additionally, our NonCog latent variable reflects the residual variance of adult educational attainment, and therefore is a measure of noncognitive aspects of adult EA. Noncognitive aspects of childhood achievement might differ somewhat, which might lead to an underestimation of indirect genetic effects of the NonCog PGS on these outcomes.

Second, the generalisability of our results is limited. Highly educated individuals are over-represented in all cohorts. Participation bias also affects GWAS results[192]. Selection effects may be especially strong in the adoption design as adoptions may depend on (partially heritable) phenotypes of the biological parents, and many adoptive parents are also selected based on their (partially heritable) behavioural phenotypes. Additionally, only participants of European descent were included in the analysis.

Third, replication efforts are needed. Special effort should be targeted to include diverse ancestry participants. While our overall estimates are well powered due to the aggregation of cohorts, some analyses rely on a single sample. As such, results from these analyses might reflect specifics of these samples and not design-specific biases and should be replicated.

Fourth, although our within-family methods allowed us to evaluate biases in polygenic score effects within the target samples, the same biases are likely to influence the effect size estimates from the original population-based GWASs used to construct polygenic scores. This problem has been explored in relation to the sibling design in a recent preprint[193], but remains to be investigated for non-transmitted PGS and adoption designs. Population

GWAS effects could be differentially affected (i.e. stronger correlation between direct and indirect genetic effects) for NonCog versus Cog, which would make their respective PGS effects less comparable. Increasingly large within-family GWASs[165,194] of Cog and EA will allow this to be resolved.

Finally, while we conceptualize our NonCog PGS as a noncognitive measure, it could also be considered a 'not-cognitive PGS', since it is a residual construct that results from removing heritable variance associated with cognitive skills from the heritable variance in educational attainment. In the future, it may be useful to develop a more precise noncognitive skills GWAS, by creating the latent Cog and NonCog factors using additional measured phenotypes. To this end, large GWA meta-analyses should be completed not only for personality[85] but not for other traditional noncognitive skills such as motivation and self-control.

Several additional future research directions emerge. First, given that we have quantified the overall environmental effects of parents on offspring education tagged by NonCog and Cog PGS, the next step is to identify specific mediating parent characteristics, whether proximal or distal. It will be informative to test not only typical noncognitive skills measures such as parental locus of control (as suggested by[146]), but also 'not-cognitive' characteristics that do not appear in noncognitive skill batteries yet are genetically correlated with the NonCog PGS and phenotypically correlated with offspring achievement. For instance, parental depression is a feasible partial mediator, given that Major Depressive Disorder is significantly genetically correlated with NonCog (r_g = -0.19, p = 2.62E^{-14})[152], and maternal depression is associated with offspring mathematics performance, possibly via offspring executive function[195]. Researchers could also examine mediating child characteristics on the pathway between their parents' characteristics and their own educational outcomes. Children's skills themselves might not be involved in these pathways. Indeed, educated parents do not appear to affect offspring education by fostering noncognitive skill development[145], and twin research shows no influence of shared environmental factors on individual differences in children's measured noncognitive skills such as grit and self-control[27,52,196].

A second future direction is to incorporate gender and socioeconomic status into research on indirect genetic effects on education. Twin data show that shared environmental contributions to educational attainment are larger for women than for men[176]. It is unknown whether this finding holds for indirect genetic effects and for childhood achievement. Another gender aspect to consider is differential maternal and paternal indirect genetic effects[163]. There is some evidence (although not genetically-informed) that mother and father skills show unique associations with offspring education[147]. Indirect effects of parents' genetic endowment for noncognitive skills on child education might be mediated or moderated by

parents' income and cultural capital (including school-related skills and habits). While some evidence suggests that home learning environments may be more cognitively stimulating in families of higher socioeconomic[197,198], there is also evidence suggesting that mothers who have lower reported incomes also report more frequent activities that facilitate cognitive stimulation[199].

In sum, this study provides evidence for environmental effects of parents' noncognitive and cognitive skills on offspring educational outcomes, indexed by indirect genetic effects of polygenic scores. Combining three cohorts and three designs for estimating indirect genetic effects allowed us to obtain robust findings. These results have significance for human health, as the role parents play in successful cognitive development and (mental) health development go hand in hand.

METHODS

Our research complies with all relevant ethical regulations. Project approval for the Twins Early Development Study (TEDS) was granted by King's College London's ethics committee for the Institute of Psychiatry, Psychology and Neuroscience PNM/09/10–104. Ethical approval for the Netherlands Twin Register (NTR) was provided by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam, and Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB-2991 under Federal-wide Assurance-3703; IRB/institute codes 94/105, 96/205, 99/068, 2003/182, 2010/359) and participants provided informed consent. The UK Biobank has received ethical approval from the National Health Service North West Centre for Research Ethics Committee (reference: 11/NW/0382). Informed consent was obtained from all human participants.

The study methods were pre-registered on the Open Science Framework (https://osf.io/mk938/) on the 24/02/2020. Additional non-preregistered analyses are indicated as such below and should be considered exploratory. Additional deviations from the pre-registration are detailed in Supplementary Note 1.

Samples

UK Biobank

The UK Biobank is an epidemiological resource including British individuals aged 40 to 70 at recruitment[200]. Genome-wide genetic data came from the full release of the UK Biobank data, and were collected and processed according to the quality control pipeline[201].

We defined three subsamples of the UK Biobank to be used for polygenic score analyses: adopted participants, a control group of non-adopted participants, and siblings. Starting with UK Biobank participants with QC genotype data and educational attainment data (N = 451,229), we first identified 6,407 unrelated adopted individuals who said yes to the question "Were you adopted as a child?" (Data-Field 1767). We restricted the sample to unrelated participants (kinship coefficient <1/(2^9/2))[202]. Second, our comparison sample (N = 6,500) was drawn at random from non-adopted participants who were unrelated to each other and to the adopted participants. Third, we identified 39,500 full siblings, excluding adopted individuals. We defined full-siblings as participants with a kinship coefficient between 1/(2^(3/2)) and 1/(2^(5/2)) and a probability of zero IBS sharing >0.0012, as suggested by[201] and [202].

After excluding the three sub-samples for polygenic score analyses and individuals related to these participants, we were left with 388,196 UK Biobank individuals with educational attainment (EA) data, and 202,815 individuals with cognitive performance (CP) data. We used these remaining individuals for the GWASs of EA and CP, and later meta-analysis with external GWASs[58] (see 'Statistical Analyses' and Supplementary Note 2).

Twins Early Development Study (TEDS)

The Twins Early Development Study (TEDS) is a multivariate, longitudinal study of >10,000 twin pairs representative of England and Wales, recruited 1994–1996[203]. The demographic characteristics of TEDS participants and their families closely match those of families in the UK. Analyses were conducted on a sub-sample of dizygotic (DZ) twin pairs with genome-wide genotyping and phenotypic data on school achievement at age 12 (1,431 DZ pairs) and age 16 (2,398 pairs).

Netherlands Twins Register (NTR)

The Netherlands Twin Register (NTR)[61] is established by the Department of Biological Psychology at the Vrije Universiteit Amsterdam and recruits children and adults twins for longitudinal research. Data on health, personality, lifestyle and others, as well as genotyping data have been collected on participants and their families.

We included in our analyses genotyped European-ancestry participants. We created a subsample of full-siblings. NTR contains information on numerous monozygotic multiples (twins or triplets). Because MZ multiples share the same genes, we randomly excluded all individuals but one per MZ multiple. Only siblings with complete genetic and outcome data were subsequently included in the analyses: 1,631 siblings with CITO (achievement test taken during the last year of primary school) data (from 757 families) and 3,163 siblings with

EA data available (from 1,309 families).

We created a subsample with complete offspring, maternal and paternal genotypic data (i.e. trios). Among individuals with available parental genotypes, respectively 1,526 (from 765 families) and 2,534 (from 1,337 families) had reported CITO and EA information.

The sibling and trio subsets are not independent: for CITO, 823 participants are present in both subsets, 1,374 for EA.

Phenotypic Measures

UK Biobank

Educational attainment and cognitive performance phenotypes were defined following Lee et al.[58]. From data-field 6,238, educational attainment was defined according to ISCED categories and coded as the number of Years of Education. The response categories are: none of the above (no qualifications) = 7 years of education; Certificate of Secondary Education (CSEs) or equivalent = 10 years; O levels/GCSEs or equivalent = 10 years; A levels/AS levels or equivalent = 13 years; other professional qualification = 15 years; National Vocational Qualification (NVQ) or Higher National Diploma (HNC) or equivalent = 19 years; college or university degree = 20 years of education. For cognitive performance, we used the (standardized) mean of the standardized scores of the fluid intelligence measure (data-field 20016 for in-person and 20191 for an online assessment).

TEDS

Educational achievement at age 12 was assessed by teacher reports, aggregated across the three core subjects (Mathematics, English, and Science).

Educational achievement at age 16 was assessed by self-reported results for standardized tests taken at the end of compulsory education in England, Wales and Northern Ireland: General Certificate of Secondary Education; GCSE). GCSE grades were coded from 4 (G; the minimum pass grade) to 11 (A*; the highest possible grade). As with the age 12 measure, we analysed a variable representing mean score for the compulsory core subjects.

NTR

Educational attainment was measured by self-report of the highest obtained degree[204]. This measure was re-coded as the number of years in education, following Okbay et al.[205].

Academic achievement is assessed in the Netherlands by a nation-wide standardized educational performance test (CITO) around the age of 12 during the last year of primary education. CITO is used to determine tracking placement in secondary school in the

Netherlands, in combination with teacher advice. The total score ranges from 500 to 550, reflecting the child's position relative to the other children taking the test this particular year.

Genotype quality control

UK Biobank

SNPs from HapMap3 CEU (1,345,801 SNPs) were filtered out of the imputed UK Biobank dataset. We then did a pre-PCA QC on unrelated individuals, and filtered out SNPs with MAF < .01 and missingness > .05, leaving 1,252,123 SNPs. After removing individuals with non-European ancestry, we repeated the SNP QC on unrelated Europeans (N = 312,927), excluding SNPs with MAF < .01, missingness >.05 and HWE $p < 10^{-10}$, leaving 1,246,531 SNPs. The HWE p-value threshold of 10^{-10} was based on: http://www.nealelab.is/blog/2019/9/17/genotyped-snps-in-uk-biobank-failing-hardy-weinberg-equilibrium-test. We then created a dataset of 1,246,531 QC-ed SNPs for 456,064 UKB subjects of European ancestry. Principal components were derived from a subset of 131,426 genotyped SNPs, pruned for LD ($r^2 > 0.2$) and long-range LD regions removed[206]. PCA was conducted on unrelated individuals using flashPCA v2[207].

TEDS

Two different genotyping platforms were used because genotyping was undertaken in two separate waves. AffymetrixGeneChip 6.0 SNP arrays were used to genotype 3,665 individuals. Additionally, 8,122 individuals (including 3,607 DZ co-twin samples) were genotyped on Illumina HumanOmniExpressExome-8v1.2 arrays. After quality control, 635,269 SNPs remained for AffymetrixGeneChip 6.0 genotypes, and 559,772 SNPs for HumanOmniExpressExome genotypes.

Genotypes from the two platforms were separately phased and imputed into the Haplotype Reference Consortium (release 1.1) through the Sanger Imputation Service before merging. Genotypes from a total of 10,346 samples (including 3,320 DZ twin pairs and 7,026 unrelated individuals) passed quality control, including 3,057 individuals genotyped on Affymetrix and 7,289 individuals genotyped on Illumina. The identity-by-descent (IBD) between individuals was < 0.05 for 99.5% in the merged sample excluding the DZ co-twins (range = 0.00 – 0.12) and ranged between 0.36 and 0.62 for the DZ twin pairs (mean = 0.49). There were 7,363,646 genotyped or well-imputed SNPs (for full genotype processing and quality control details, see[208]).

To ease high computational demands for the current study, we excluded SNPs with MAF <1% and info < 1. Following this, 619216 SNPs were included in polygenic score construction.

Principal components were derived from a subset of 39,353 common (MAF > 5%), perfectly imputed (info = 1) autosomal SNPs, after stringent pruning to remove markers in linkage disequilibrium (r2 > 0.1) and excluding high linkage disequilibrium genomic regions to ensure that only genome-wide effects were detected.

NTR

Genotyping was done on multiple platforms, following manufacturers protocols: Perlegen-Affymetrix, Affymetrix 6.0, Affymetrix Axiom, Illumina Human Quad Bead 660, Illumina Omni 1M and Illumina GSA. For each genotype platform, samples were removed if DNA sex did not match the expected phenotype, if the PLINK heterozygosity F statistic was < -0.10 or > 0.10, or if the genotyping call rate was < 0.90. SNPs were excluded if the MAF < 1×10^{-6}, if the Hardy-Weinberg equilibrium p-value was < 1×10-6, and/or if the call rate was < 0.95. The genotype data was then aligned with the 1000 Genomes reference panel using the HRC and 1000 Genomes checking tool, testing and filtering for SNPs with allele frequency differences larger than 0.20 as compared to the CEU population, palindromic SNPs and DNA strand issues. The data of the different platforms was then merged into a single dataset, and one platform was chosen for each individual. Based on the ~10.8k SNPs that all platforms have in common, DNA identity-by-descent state was estimated for all individual pairs using the Plink 1.9 and King 2.1.6 programs. Samples were excluded if these estimates did not correspond to expected familial relationships. CEU population outliers, based on per platform 1000 Genomes PC projection with the Smartpca software v2.r904, were removed from the data. Then, per platform, the data was phased using Eagle v2.4.1 and then imputed to 1000 Genomes and Topmed using Minimac3-omp v2.10 following the Michigan imputation server protocols. Post-imputation, the resulting separate platform VCF files were merged with Bcftools 1.9 into a single file per chromosome for each reference, for SNPs present on all platforms. For the polygenic scoring and parental re-phasing, the imputed data were converted to best guess data and were filtered to include only ACGT SNPs, SNPs with MAF > 0.01, HWE p > 10 -5 and a genotype call rate > 0.98, and to exclude SNPs with more than 2 alleles. All mendelian errors were set to missing. The remaining SNPs represent the transmitted alleles dataset. 20 PCs were calculated with Smartpca using LD-pruned 1000 Genomes–imputed SNPs genotyped on at least one platform, having MAF > 0.05 and not present in the long-range LD regions. Using the --tucc option of the Plink 1.07 software pseudo-controls for each offspring were created, given the genotype data of their parents. This resulted in the non-transmitted alleles dataset, as these pseudo-controls correspond to the child's non-transmitted alleles. To determine the parental origin of each allele, the transmitted and non-transmitted datasets were phased using the duoHMM option of the

ShapeIT software. The phased datasets were then split based on parental origin, resulting in a paternal and maternal haploid dataset for the transmitted and non-transmitted alleles.

Statistical analyses

All statistical tests are two-sided, unless otherwise stated.

NonCog GWAS-by-subtraction

To generate NonCog summary statistics, we implemented a GWAS-by-subtraction using Genomic-SEM following Demange et al. 2020 using summary statistics of EA and cognitive performance obtained in samples independent from our polygenic score samples.

We ran a GWAS of Educational Attainment and Cognitive Performance in UK Biobank (polygenic score sample left-out). We meta-analysed them with the EA GWAS by Lee et al. excluding 23andMe, UK Biobank and NTR cohorts, and with the CP GWAS by Trampush et al. respectively (EA total N = 707,112 and CP N = 238,113) using Metal software release 2011-03-05. More information on these methods and intermediate GWASs are found in Supplementary Note 2 and Supplementary Data 1.

Following Demange et al. 2020, we used EA and CP meta-analysed summary statistics with Genomic-SEM to create two independent latent variables: Cog, representing the genetic variance shared between EA and CP, and NonCog representing the residual genetic variance of EA when regressing out CP (Supplementary Figure 1). These two latent factors were regressed on each SNP: we obtained association for 1,071,804 SNPs (HapMap3 SNPs, as recommended when comparing PGS analyses across cohorts). We calculate the effective sample size of these GWASs to be 458,211 for NonCog and 223,819 for Cog.

Polygenic Score construction in UK Biobank, TEDS and NTR

Polygenic scores of NonCog and Cog were computed with Plink software (version 1.9 for NTR, 2 for UKB and TEDS)[209,210] based on weighted betas obtained using the LDpred v1.0.0 software using infinitesimal prior, a LD pruning window of 250kb and 1000Genomes phase 3 CEU population as LD reference. Weighted betas were computed in a shared pipeline. In NTR, scores for non-transmitted and transmitted genotypes were obtained for fathers and mothers separately so we average them to obtain the mid-parent score.

Polygenic score model fitting

Each model included cognitive and noncognitive polygenic scores simultaneously and controlled for: 10 ancestry principal components (PCs), sex and age, interaction between sex and age, and cohort-specific platform covariate (NTR: genotyping platform, UKB: array, TEDS: batch). Age was estimated by year of birth, age at recruitment or age at testing

depending on the cohorts, see Supplementary Data 2. Correlations between NonCog and Cog PGS, as well as between and within-family PGS are reported Supplementary Data 10.

Outcomes were standardized for each analysis group. Polygenic scores were standardised as follows prior to analysis. For the non-transmitted allele design, we summed the parental PGS and then scaled the non-transmitted and transmitted PGS separately, following Kong et al[156]. Note that the variances of the non-transmitted and transmitted PGS were not significantly different prior to scaling (Cog PGS: $F = 1.0088$, $p = 0.71$; NonCog PGS: $F = 0.9920$, $p = 0.73$). For the adoption design, we scaled the PGS in adopted and non-adopted groups separately. There were no significant differences in variances of adopted and non-adopted PGS prior to scaling (see Supplementary Data 11). For the sibling design, we scaled the PGS to have mean 0 SD 1 using the sibling group, and subsequently created the within-sibling PGS.

All regressions were linear models with lm() in R rather than mixed models as in previous analyses[157,158] and our pre-registered methods. See Supplementary Note 1 for the justification based on simulated data. We obtained bootstrapped standard errors and bias-corrected confidence intervals (normal approximation) for the population, indirect and direct effects, as well as the ratios of indirect/direct and indirect/population effect. We ran ordinary non-parametric bootstraps using 10,000 replications with boot() in R. For the sibling design, where two independent regressions are used, we use the same bootstrap samples for both (both regressions were run within the same boot object). For the adoption design, the bootstrapped samples are drawn from the adopted and non-adopted samples separately. The bootstrap estimates were used to test for the difference between the direct and indirect effect in both Cog and NonCog and the difference between the ratio indirect/population for Cog and NonCog, using Z-tests.

Additional analyses (not pre-registered)

Meta-analyses

To estimate the overall indirect and direct effects of NonCog and Cog polygenic scores, we meta-analysed estimates across cohorts, designs and phenotypic outcomes.

To compare results obtained across the three different designs, we meta-analysed effect sizes obtained from each design across cohorts, but holding the outcome constant (educational attainment). The adoption design was only applied to EA in UKB, hence no meta-analysis was necessary.

Meta-analyses were conducted using the command rma.mv() in the R package metafor. Design was specified as a random intercept factor, except when results were meta-analysed

within-design.

Investigation of biases

Population stratification

Population stratification refers to the presence of systematic difference in allele frequencies across subpopulations, arising from ancestry difference due to non-random mating and genetic drift. This leads to confounding in genetic association studies. In a PGS analysis, bias due to population stratification can arise from both the GWAS used to create the scores and the target sample. We corrected for population stratification in the target sample by adjusting analyses for PCs (although this may not remove fine-scale stratification). For the GWAS summary statistics, the ratio statistics LDSC output is a standard measure of population stratification[211]. As a rule of thumb, an LDSC intercept higher than 1 (inflated) indicates presence of population stratification. Because we corrected the standard errors of the EA GWAS for inflation and Genomic-SEM corrects for inflation as well, the ratio statistics of the Cog and NonCog GWASs are not a valid indication of population stratification (ratio <0 following GC correction). We therefore use the ratio statistics of uncorrected EA and CP GWASs as proxies. Ratio and LDscore intercept was assessed with the ldsc software[211].

Assortative mating

Assortative mating refers to the non-random mate choice, with a preference for spouses with similar phenotypes. If these preferred phenotypes have a genetic component, assortative mating leads to an increased genetic correlation between spouses, as well as between relatives[181]. Assortative mating can therefore be inferred from elevated correlations between polygenic scores in siblings (correlations would be 0.5 without assortative mating) and between parents (correlations would be 0 without assortative mating). We estimated sibling intraclass correlations of Cog and NonCog PGS in UKB, TEDS and NTR, and Pearson's correlations of paternal and maternal Cog and NonCog PGS in NTR. Notably, these observed correlations cannot distinguish assortative mating from population stratification.

Sibling effects

We performed three additional analyses to investigate indirect genetic effects of siblings on educational outcomes.

First, we ran a linear mixed model extending our main non-transmitted alleles design to include polygenic scores of siblings (Eq. (4)). To this end, we used data from NTR on DZ pairs and both of their parents. Sample sizes of genotyped 'quads' with offspring CITO or EA phenotypes were 657 and 788, respectively.

$$\begin{aligned} EA = {} & \alpha_0 0 + \beta_{T_{Cog}}(PGS(Cog)_T) + \beta_{T_{NonCog}}(PGS(NonCog)_T) \\ & + \beta_{NT_{Cog}}(PGS(Cog)_{NT}) + \beta_{NT_{NonCog}}(PGS(NonCog)_{NT}) \\ & + \beta_{Sibling_{Cog}}\left(PGS(Cog)_{Sibling}\right) \\ & + \beta_{Sibling_{NonCog}}\left(PGS(NonCog)_{Sibling}\right) + sex + age + sex \\ & * age + 10PCs + genotyping\ platform \end{aligned} \quad (4)$$

Second, we can also assess the presence of sibling genetic effects using monozygotic and dizygotic twins. Because monozygotic twins have the same genotypes, the genetically mediated environment provided by the cotwin is more correlated to the twin genotype in MZ twins than in DZ twins. The sibling genetic effect is more strongly reflected in the polygenic score prediction of the educational outcome for MZ twins than for DZ twins. If the sibling genetic effect is negative, the polygenic score effect (betas) on the outcome in people that have an MZ twin will be lower than in people that have a DZ twin, it will be higher in those with an MZ twin then those with an DZ twin if the sibling genetic effect is positive. We therefore compare Betas from Eq. (2) in a subset of MZ twins and in a subset of DZ twins (one individual per pair) in both NTR (N_{MZ} = 818 & N_{DZ} = 865 for CITO and N_{MZ} = 1,600 & N_{DZ} = 1,369 for EA) and TEDS (N_{MZ} = 546 & N_{DZ} = 2,709)

Third, the presence of sibling genetic effects can be assessed using data on the number of siblings participants have. If an individual has more siblings, we expect their polygenic scores to be more correlated to sibling effects. As the number of siblings increases (if we assume linear increase) so does the degree to which a PGS captures sibling effects. If the sibling genetic effect is positive, the effect of the Cog and NonCog PGS on the educational outcome should increase with the number of siblings. However, family characteristics and environment might differ across families depending on the number of children. Therefore, changes in the effect of the PGS on our outcome with the number of siblings could be due to factors other than sibling genetic effects (for example, there is a known negative genetic association between number of children and EA[84] which could result in confounding). By also looking at changes in the effect of the Cog and NonCog PGS on the educational outcome in adopted (unrelated) sibships, we break the correlation between PGS and any sibling effects. If there is a change in PGS effect on the educational outcome in adopted children dependent on the number of (non-biological) siblings, we can assume this effect to be caused by mechanisms other than a sibling effect. We finally contrast the change in PGS depending on family size in biological and adopted siblings to get an idea of the sibling effect minus any other confounding effects of family size. We use the total number of reported siblings (full siblings for non-adopted and adopted siblings for adopted individuals, data-

fields: 1873, 1883, 3972 & 3982).

CODE AND DATA AVAILABILITY

All scripts used to run the analyses (empirical and simulated) are available at our GitHub https://github.com/PerlineDemange/GeneticNurtureNonCog/, which can be cited as Demange P., et al. Estimating effects of parents' cognitive and noncognitive skills on offspring education using polygenic scores, GitHub, DOI: 10.5281/zenodo.6581326, 2022.

All additional software used to perform the analyses are available online.

The pre-registration of the study is available on OSF: https://osf.io/mk938/

For the original summary statistics of Cog and NonCog, including NTR and UK Biobank siblings data, see152. The summary statistics for Cog and NonCog generated for this study are available at: https://doi.org/10.34894/MMXYPL.

For UK Biobank dataset access, see: https://www.ukbiobank.ac.uk/using-the-resource/.

Netherlands Twin Register data may be accessed, upon approval of the data access committee, email: ntr.datamanagement.fgb@vu.nl

Researchers can apply for access to TEDS data: https://www.teds.ac.uk/researchers/teds-data-access-policy

SELECTED SUPPLEMENTARY INFORMATION

Full Supplementary Information and Supplementary Tables can be downloaded at: https://www.nature.com/articles/s41467-022-32003-x#Sec32

Supplementary Note 4: Methods for simulating genetic and phenotypic data in the presence of different biases and components

We simulate data introducing various potential components and biases, and then fit all models used throughout the paper to identify how the estimated parental indirect effect changes in the presence of these factors.

We simulate genotype data for 20,000 families. Each family includes a mother, a father, a focal offspring, a child sibling, and an adopted child sibling. The adoptee genotypes are drawn from another simulated dataset of biological parents, independent of the focal families. Therefore, the total sample size including the main families plus biological parents of adoptees is (20,000 x 5) + 20,000 = 120,000 individuals. Genotypes are simulated as 100 bi-allelic SNP calls, using the 'coin flipping' function in R *rbinom*(). For individuals in the parent generation, probability values for SNPs are defined by minor allele frequencies (simulated as deviates of the uniform distribution between .1 and .5). For offspring of these

individuals, probability values for SNPs are defined as each parental genotype divided by 2.

We then simulate 'true' SNP effects, drawn from a normal distribution. We use these true SNP effects to simulate 'true' genetic scores for the mothers, fathers, and biological and adopted offspring. The true SNP effects are the same for all individuals and for all sub-populations. True genetic scores are used to simulate phenotypes.

In addition to 'true' polygenic scores, we create more realistic 'GWAS-based' polygenic scores for all individuals by weighting their genotypes by GWAS SNP effects. We define GWAS SNP effects as true SNP effects with added error, and calculate them as sqrt(.2)*true effects + sqrt(.8)*error, the error following a normal distribution (this differs when simulating population stratification, see below). GWAS effects are the same for all individuals and sub-populations. GWAS-based polygenic scores are used to estimate direct and indirect effects. We also tested how sensitive the estimates from the three designs are to the amount of noise introduced in the GWAS effects, and found that this only matters for assortative mating (see assortative mating results below).

We simulate nine offspring phenotypes influenced by different factors:

i. direct genetic effects only,
ii. direct and indirect parental genetic effects (maternal and paternal),
iii. indirect parental genetic effects plus a prenatal indirect maternal genetic effect,
iv. indirect sibling genetic effect,
v. indirect parental genetic effects and an indirect sibling genetic effect,
vi. assortative mating,
vii. assortative mating and indirect parental genetic effects,
viii. population stratification,
ix. population stratification and indirect parental genetic effects.

Having simulated the nine phenotypes as detailed further below, we use three designs (sibling, adoption, non-transmitted allele, explained in the main article) to estimate indirect parental genetic effects on each phenotype. This allows us to evaluate how designs are affected by the components (prenatal and postnatal parental indirect genetic effects) and biases (sibling indirect genetic effects, assortative mating, population stratification). We repeated the simulation 100 times.

Note that these simulations are to illustrate how designs are affected by the components and biases. Effect sizes for each bias/component are not intended to represent true effects and as such are somewhat arbitrary. Additionally, by necessity we make certain untested

assumptions. For example, indirect genetic effects are assumed to be equal between all siblings (i.e. no birth order effects or different effects for adoptive siblings), and population stratification and assortative mating are assumed to operate equally among biological and adoptive parents.

For the simulation results, see the below text under the heading 'Comparison of sibling, adoption, and non-transmitted allele designs in presence of simulated components and biases', and Supplementary Figure 9. The complete simulation code is available on GitHub[169].

Simulation details for the nine phenotypes

i. Direct genetic effects

We simulate child phenotypes influenced by direct genetic effects only, such that

$$y = var(g).x + e$$

Where y is the child phenotype, x is the true genetic score of the child, var(g) is the variance explained by the true genetic score and e is the residual error (explaining the rest of the variance).

Parental phenotypes used below are also simulated this way (i.e. influenced by own genotype plus environment/error).

ii. Indirect parental genetic effects

We simulate child phenotypes influenced by direct genetic effects and indirect parental genetic effects such that

$$y = var(g).x + var(mother).y_{mother} + var(father).y_{father} + e$$

Where y is the child phenotype, x is the true genetic score of the child, var(g) is the variance explained by the true genetic effect, y_{mother} and y_{father} are the parental phenotypes, *var(mother)* and *var(father)* are the variance explained by parental phenotypes, and e is the residual error (explaining the rest of the variance).

iii. Prenatal and postnatal indirect parental effects

We simulate child phenotypes influenced by direct genetic effects and prenatal and postnatal indirect parental genetic effects such that

$$y_{adoptee} = var(g).x + var(prenatal).y_{biological\ mother} + var(postnatal).y_{adoptive\ mother} + var(father).y_{father} + e$$

$$y_{non-adopted} = var(g).x + var(prenatal).y_{mother} + var(postnatal).y_{mother} + var(father).y_{father} + e$$

iv. Indirect sibling genetic effects &

v. Indirect sibling and parental indirect genetic effects

After simulating all sibling phenotypes with only direct effects or with direct and indirect parental genetic effects, we simulate indirect genetic effects operating among three siblings in each family: individual 1, a biological sibling, and an unrelated adopted sibling. First, we create a matrix of sibling effects in which every effect is of the same magnitude (all siblings have an equal effect on each other regardless of adoption status, an implicit assumption), with zeros on the diagonal. To account for feedback effects (e.g. sibling 1 influences sibling 2, who influences sibling 1; this changes the coefficients of a variable on its own errors), we subtract the sibling effect matrix from an identity matrix and take its inverse. We then take the matrix product of the matrix with sibling effects and the simulated sibling data to introduce the simulated mutual sibling effects into the data.

vi. Assortative mating

vii. Assortative mating and parental indirect genetic effects

Genetic assortative mating occurs when individuals with similar phenotypes mate more frequently than would be expected under a random mating scenario, and these phenotypes are heritable. To simulate assortment, we re-create offspring genotypes and polygenic scores after matching parents together systematically (instead of randomly as above). We first create phenotypes for the parents (based on true genetic score plus noise), rank the mothers and fathers by phenotype, and match couples according to rank (i.e. mothers with higher phenotypic values match with fathers with higher phenotypic values). Since mating does not perfectly track with phenotypic rank, we add noise to the ranking of mothers and fathers prior to matching, following a chosen phenotypic correlation. Offspring genotypes are then simulated as random draws from the matched couples' genotypes. Assortment is simulated to be the same strength for adoptees' and nonadoptees' parents, and we simulate random placement by un-ranking adoptees before matching them to adoptive families.

viii. Population stratification

ix. Population stratification and parental indirect genetic effects

Population stratification can be conceptualized as systematic differences in allele frequencies between sub-populations. These frequency differences cause confounding in genetic studies when phenotypes also differ between sub-populations. We simulate such sub-populations in both the GWAS discovery and target PGS analyses samples. We first create new genotypes in two groups, drawing upon two different sets of simulated minor allele frequency distributions. We also define a phenotypic difference between these

two groups, by including an 'environmental confounding' parameter which is noise with a different mean phenotype for the two sub-populations. We then run a single GWAS in these two populations. We create phenotypes and polygenic scores (based on the GWAS results) in a target sample of families, comprising the same two sub-populations present in the GWAS. Our simulation allows for adoptees to be matched with adoptive parents both within- and between- sub-populations. We report results from a simulation with adoptees matched with adoptive parents within the same sub-population.

Supplementary Note 5: Comparison of two implementations of the sibling design using simulation

In the sibling design presented by Selzam et al.[158], indirect genetic effects are estimated by subtracting the within-sibling estimate from the between-sibling estimate (indexed using the average polygenic score for each sibling pair). However, the between-sibling effect is not necessarily the appropriate quantity to use[212]. An alternative is to subtract the within-sibling estimate from an estimate of the population effect obtained in a separate regression analysis using population data and ignoring family clustering. This approach was used in a recent within-sibling GWA study[194].

To ensure that we contrast our direct genetic effects with the appropriate quantity for accurate estimation of indirect genetic effects, we use simulated data to assess the use of the between-sibling effect and the population effect. Results are presented below in Supplementary Figure 8. From these simulations, it appears that contrasting the direct effects with the between-sibling effects leads to an overestimation of indirect parental genetic effects. Contrasting direct effects with population effects results in accurate estimation of indirect genetic effects. Consequently, we use this approach in our main analyses and simulations. Therefore, our model differs slightly from the Selzam et al. analyses.

Also notable is that, whilst the Selzam et al. article (and[159]) uses a different term – passive gene-environment correlation – the effect being estimated is a parental indirect genetic effect. Passive gene-environment correlation refers to how the genes that parents pass on to their children may also influence how they provide the rearing environment[213].

Supplementary Note 6: Comparison of sibling, adoption, and non-transmitted allele designs in presence of simulated components and biases

Using the simulated data, we compare the behaviour of the three designs used in our study to estimate direct and indirect genetic effects. For simplicity's sake our simulations consider one PGS (instead of both *Cog* and *NonCog* PGS). Additionally, we compare to a fourth design which we call "trios" in Supplementary Figure 9, in which the phenotype is

simply regressed on child and parental PGS. As the simulation results prove, this simple approach gives identical estimates to the non-transmitted allele design, which also uses trios but requires prior identification of segments that are shared and non-shared between the generations.

Supplementary Figure 9 (an extended version of Figure 3 in the main text) displays the simulation results. The following text discusses the results, focusing on the main estimates of interest – indirect genetic effects of parents.

Prenatal parental indirect genetic effects

We see prenatal effects as a component of interest, rather than as a bias, in estimates of indirect genetic effects. Nonetheless, for consistency with the rest of the simulations which do not consider prenatal indirect genetic effects, the red dashed line in Supplementary Figure 9 indicates the true postnatal effect only. Simulation results show that the sibling- and trio-based designs capture indirect genetic effects occurring in both prenatal and postnatal periods. In contrast, the adoption design only captures postnatal indirect genetic effects. This is because, for both adoptees and non-adopted individuals, the prenatal environment is provided by the biological mother, so estimated polygenic score-phenotype associations for both adoptees and non-adopted individuals contain prenatal maternal indirect genetic effects. Consequently, computing the indirect genetic effect as the population effect of the polygenic score (β in non-adopted individuals) minus the direct genetic effect (β for adoptees) means that prenatal effects are cancelled out. This result suggests that prenatal indirect genetic effects could partially explain lower estimates of indirect genetic effects from the adoption design compared to the other designs.

Sibling indirect genetic effects

We find that positive sibling effects result in upwardly biased estimates of indirect parental genetic effects. This bias is considerably larger for the sibling design than the adoption and trio designs. Bias in the sibling design is likely to be because positive sibling effects increase the similarity of siblings, reducing the effect of within-sibling polygenic differences[214,215]. Bias in the non-transmitted allele design likely arises because non-transmitted alleles are not only shared with parents, but also partially with (full) siblings, such that βNT might capture sibling as well as parental indirect genetic effects. It is interesting that sibling effects inflate adoption-based estimates despite the fact that adoptees are not genetically related to their siblings. These simulation results lead to the notion that higher estimates of parental indirect genetic effects in the sibling than adoption and non-transmitted allele designs is evidence of sibling indirect genetic effects. In our empirical data, we do not find differences

between sibling- and trio-based estimates of indirect parental genetic effects. Along with our sensitivity analyses, this suggests an absence of sibling genetic effects on educational outcomes in our datasets.

Assortative mating

In our main scenario, which includes substantial error in the GWAS SNP effects used to calculate polygenic scores, so the correlation of GWAS SNP effects and true SNP effects is on average 0.45), we found that the bias from assortative mating in the indirect genetic effect estimate was lower in the adoption design than in the non-transmitted allele and sibling designs. We also tested other scenarios with lower error in the SNP effects used to make the polygenic score. In the scenario with assortative mating but not indirect effects, lower error in the polygenic scores led to decreased bias in estimates from the NT and sibling designs. In the scenario with both assortative mating and indirect effects, with decreasing error in SNP effects, the sibling estimate is consistently biased, but the adoption estimate is more biased and the NT estimate is less biased. Results of other simulations did not change according to the error. We present in the main manuscript the initial results with substantial error as the most conservative example. In real data, we expect this bias due to the combination of error in effect sizes and non-random mating to decrease as GWAS sample sizes increases.

Bias in the sibling design likely arises as the population effect contains assortative mating while the within-sibling effect does not. Bias in the non-transmitted allele design due to assortative mating, which happens due to correlations between parental alleles, is described in Kong et al. 2018[156]. Interestingly, the bias in the adoption design from assortative mating is zero in the absence of a parental indirect genetic effect, but slightly above zero when a parental indirect genetic effect was also specified. In other words, the presence of parental indirect genetic effects is required for assortative mating to bias estimates from the adoption design. We simulated the same strength of assortative mating for the parents of both adopted and non-adopted individuals, so the result cannot be due to elevated assortment in the latter group (leading to residual assortment in the indirect effect estimate when calculating $\beta_{non\text{-}adopted} - \beta_{adopted}$). Such differences could exist in the real data, but there is scarce and inconsistent evidence regarding assortment in biological parents of adoptees versus other parents[191,213]. Overall, the results suggest that assortative mating could explain lower estimates of indirect genetic effects from the adoption design compared to the other designs, but may depend on the level of noise in the GWAS effects.

Population stratification

Simulation results show that estimates of parental indirect genetic effects based on the

adoption design capture less bias from population stratification than sibling- and trio-based designs. In the sibling design, the parental indirect genetic effect is estimated as the population effect minus the direct within-family effect of the polygenic score. This means that the indirect genetic effect is likely to be inflated by population stratification, as this is captured in the population effect but not the within-family effect. Also, the effects of the non-transmitted allele PGS are influenced by population stratification, so the indirect genetic effect estimate is inflated. In contrast, population stratification only biases indirect genetic effect estimates from the adoption design to a small extent. Assuming that population stratification is similar in adoptees and non-adopted individuals, its effect will cancel out when estimating the indirect genetic effect as $\beta_{non\text{-}adopted} - \beta_{adopted}$. The assumption of equal population stratification and assortative mating bias in adopted and non-adopted groups cannot be tested due to the lack of parental data in UKB, but is bolstered by the simulation results, and by the fact that both adoptees and non-adopted individuals are from British ancestry. Our simulation results suggest that population stratification partly explains the lower estimates of indirect genetic effects from the adoption design compared to the other designs in our empirical study.

CHAPTER 4

EVALUATING THE CAUSAL RELATIONSHIP BETWEEN EDUCATIONAL ATTAINMENT AND MENTAL HEALTH

Under review and preprinted as: **Demange, P.A.**, Boomsma, D.I., van Bergen, E.*, Nivard, M.G.* (2023) Evaluating the causal relationship between educational attainment and mental health. medRxiv, 2023.01.26. https://doi.org/10.1101/2023.01.26.23285029

Supplementary materials accessible at:

https://www.medrxiv.org/content/10.1101/2023.01.26.23285029v1.supplementary-material

ABSTRACT

We investigate the causal relationship between educational attainment (EA) and mental health using two research designs. First, we compare the relationship between EA and seventeen psychiatric diagnoses within sibship in Dutch national registry data (N = 1.7 million), controlling for unmeasured familial factors. Second, we use two-sample Mendelian Randomization, which uses genetic variants related to EA or psychiatric diagnosis as instrumental variables, to test whether there is a causal relation in either direction. Our results suggest that lower levels of EA causally increase the risk of MDD, ADHD, alcohol dependence, GAD and PTSD diagnoses. We also find evidence of a causal effect in the opposite direction for ADHD. Additionally, we find inconsistent results for schizophrenia, anorexia nervosa, OCD, and bipolar disorder, highlighting the importance of using multiple research designs to understand the complex relationship between EA and mental health.

INTRODUCTION

Over 17% of the population of the European Union is diagnosed with a mental disorder (2016 estimates, OECD, 2018). Mental diagnoses account for 25% of years lived with disability worldwide[217]. The risk of being diagnosed with a mental disorder is higher among those with lower educational attainment (EA)[36,218,219]. If the relationship between education and mental diagnoses is indeed causal, educational policies primarily aimed at improving educational outcomes could also lead to improved mental health.

Most prior studies of the relationship between education and mental health are correlational. This correlation could therefore reflect confounding factors influencing both education and mental health. While known and perfectly measured confounders can be controlled, unknown and unmeasured confounders cannot. Additionally, the correlation between EA and mental health might also be explained by reverse causation, as early onset of mental disorders may hamper subsequent school attendance and performance[220]. Randomized experiments in which education is altered would avoid bias due to confounding or reverse causation, but experiments at the required scale are not feasible for practical and ethical reasons.

As an alternative, we rely on two natural experiments that account for confounding and/or reverse causation in different manners: within-sibship regression and mendelian randomization (MR).

In the Dutch population registry (N = 1.7 million siblings born between 1965 and 1985), we test whether differences in EA between siblings relate to differences in their risk of psychiatric diagnoses between 2011 and 2016. A core assumption of the within-sibship design is that siblings constitute a well-matched case-control group[212,221]. Siblings are comparable for many factors that might play a role in both EA and mental health, for example the family, school, and neighbourhood environment, and 50% of their segregating genome. By comparing siblings, we obtain estimates of the association of education with psychiatric diagnoses and care expenditures, controlled for unmeasured confounders shared by the siblings. While within-sibship estimates can increase or decrease our confidence in the presence of a causal relation, they are insufficient. Confounders not shared between siblings such as differential experiences, 50% of the segregating genome, but also measurement error might bias the estimate[222]. Within-sibship association does not offer evidence of a causal direction, and while timing of events may suggest direction, unmeasured prodromal signs and experiences anterior to graduation might affect the association.

To mitigate the uncertainties introduced by limitations of the within-sibship design, we also apply mendelian randomization (MR). In MR, genetic variants that are robustly associated with the exposure are used as instrumental variables. MR's core assumptions are: (1) some

variants are associated with the exposure (EA), (2) these variants are related to mental health only via their effect on educational success, and (3) these variants do not correlate with any confounders of the relationship between education and mental health[223]. When assumptions are met, MR estimates the causal effect of education on mental health, even in presence of confounding and measurement error[224]. Reverse causation can be empirically evaluated by running two sets of MR analyses: one with variants related to EA as exposure and variants related to psychiatric diagnoses as outcome, and the reverse analysis. We apply two-sample MR[225], which uses genetic effect estimates from existing well-powered genome-wide association studies (GWASs) of EA[21] and of mental disorders in European-ancestry samples. To minimize the influence of pleiotropy (i.e. one genetic variant affects many traits), we used additional weak-instrument- and pleiotropy-robust MR methods[226–228]. To partly mitigate the influence of assortative mating, population stratification, and gene-environment correlation we performed MR based on genetic associations with EA obtained in a within-sibship GWAS[229]. Mendelian randomization applied to discrete or (ordered) categorical traits, such as EA and mental disorders, has an additional notable caveat: interpretation[230]. If we assume that genetic variants influence categorical variables via their effects on the underlying liability, MR estimates the effect of the liability for higher EA and the liability for being diagnosed, while the within-sibship design estimates the effect of the observed exposure on the risk of the observed outcome.

Only a few prior quasi-causal experiments investigate EA and mental disorders, and found mixed evidence. Studies that compared monozygotic twin pairs with discordant educational outcomes found evidence consistent with a causal association between EA and depressive symptoms[231], while others did not[232,233]. Similarly, MR studies showed mixed evidence for a negative effect of EA on the risk for depression diagnoses[234–236] but no reverse effect[237]. MR studies suggested a negative effect of EA on ADHD, no effect on PTSD and schizophrenia, but also a positive effect on schizophrenia, bipolar disorder, anorexia, autism and anxiety[238–240]. Within-sibship studies of mental disorders and EA are rare, and even rarer in population registries[241,242].

The reliance on quasi-causal methods, though preferred over observational association, is still imperfect, and calls for epistemic humility when interpreting or generalizing putative causal association. We reduce uncertainty on the nature of the relation between education and mental health by triangulating across two preregistered quasi-causal methods relying on different underlying assumptions, whose violation would give rise to different biases[243]. Consistent findings strengthen our confidence. Inconsistent findings raise scepticism yet are also valuable: we interpret these differences in the light of the different assumptions of

each method to help us hypothesize on the mechanisms in the relation between education and mental diagnoses.

RESULTS

In the following, we focus on diagnoses for which GWASs are available. Results for all diagnoses described in Table 1 are available in Supplementary Note and Tables.

Descriptive Analysis

In the Dutch population register, we selected siblings (sharing the same legal mother and father) born between 1965 and 1985, such that they are expected to have obtained their highest diploma before the first year of diagnostic data is available in 2011. We obtained a final sample of 1,743,032 siblings nested within 766,514 families (Supplementary Note and Supplementary Figure 1).

Inferring the number of years of education from the final degree obtained, Dutch siblings attend education for an average of 15.35 years (SD = 2.80, median = 17). The sibling sample appears slightly more educated than all individuals born between 1965-1985, see Supplementary Note and Supplementary Table 2.

We accessed psychiatric diagnoses based on the Diagnostic and Statistical Manual of Mental Disorders 4th Education (DSM-IV) for all patients getting specialized mental care in the Netherlands between 2011 and 2016. Between these years, the yearly incidence of psychiatric diagnoses decreased (Supplementary Tables 3-4). This may reflect changes in access to specialized mental health care, e.g. a 2014-reform led to an increase in the care of chronic mental health disorders by general practitioners.

Individuals can be diagnosed with more than one disorder within a year, with one primary diagnosis and one or more secondary diagnoses (Supplementary Table 7), or across the 6 years we studied. Psychiatric diagnoses co-occur frequently (Supplementary Table 6 and Supplementary Figure 2) and are highly correlated (polychoric correlation up to 0.55 for MDD and PTSD). Schizophrenia is one exception as it is only weakly correlated with most other diagnoses (Figure 1, Supplementary Table 6 & Supplementary Figure 2). Genetic correlations across the GWASs selected for the MR analyses are also positive and generally substantial, but they do not always match phenotypic correlations in the Dutch population.

		Subsamples of individuals in CBS born between 1965 and 1985				
		All (1)	All with EA (2)	Siblings (3)	Siblings with EA (4)	
	Total sample size	6,539,767	3,305,733	3,234,923	1,743,032	
	EA average (SD)	–	14.7 (3.4)	–	15.4 (2.8)	
Percentage of the population diagnosed between 2011 and 2016	Individuals without any disorder	93.00	89.06	91.24	89.43	
	Individuals with at least one disorder	7.00	10.94	8.76	10.57	**GWAS**
	Attention deficit hyperactivity disorder (ADHD)	0.93	1.53	1.31	1.62	Demontis
	Autism spectrum disorder (ASD)	–	–	–	–	Grove
	Alcohol dependence	0.92	1.41	1.12	1.30	Walters
	Schizophrenia/schizophreniform/ schizoaffective disorders	0.39	0.52	0.48	0.51	–
	Schizophrenia	0.31	0.42	0.38	0.41	Trubetskoy
	Major depressive disorder (MDD)	2.81	4.42	3.34	4.08	Howard
	Bipolar disorder	0.22	0.33	0.29	0.34	Mullins
	Bipolar I	0.15	0.22	0.19	0.23	Mullins
	Bipolar II	0.08	0.13	0.11	0.14	Mullins
	Anxiety disorders	2.50	3.94	3.01	3.70	–
	Generalized anxiety (GAD)	0.42	0.67	0.56	0.70	Purves
	Panic	0.64	1.00	0.81	0.97	–
	Phobia	0.69	1.10	0.91	1.12	–
	Obsessive-compulsive (OCD)	0.28	0.42	0.38	0.45	Arnold
	Post-traumatic stress (PTSD)	1.16	1.84	1.21	1.51	Nievergelt
	Anorexia nervosa	0.03	0.05	0.05	0.06	Watson
	Bulimia nervosa	0.05	0.08	0.06	0.08	–
	Personality disorders	3.00	4.77	3.91	4.81	–
	Cluster A	0.09	0.14	0.11	0.14	–
	Cluster B	0.95	1.55	1.19	1.47	–
	Cluster C	2.04	3.26	2.75	3.38	–

Table 1 | Prevalence of diagnoses in the Dutch population registry (CBS) between 2011 and 2016 and source of GWASs for matching diagnoses. The frequency of DSM-IV Diagnoses in the Dutch population registry registered in second-line/specialized mental health care, both as primary or secondary diagnoses. The definition of the disorder in the equivalent GWAS does not always perfectly align with the diagnosis in CBS, see Supplementary Table 18 for details of GWASs used. Subsamples are relevant subsets of the Dutch population registry that met criteria for inclusion in the study. Inclusion is further described in the method section and the chart Supplementary Figure 1. EA: educational attainment. SD: Standard deviation.

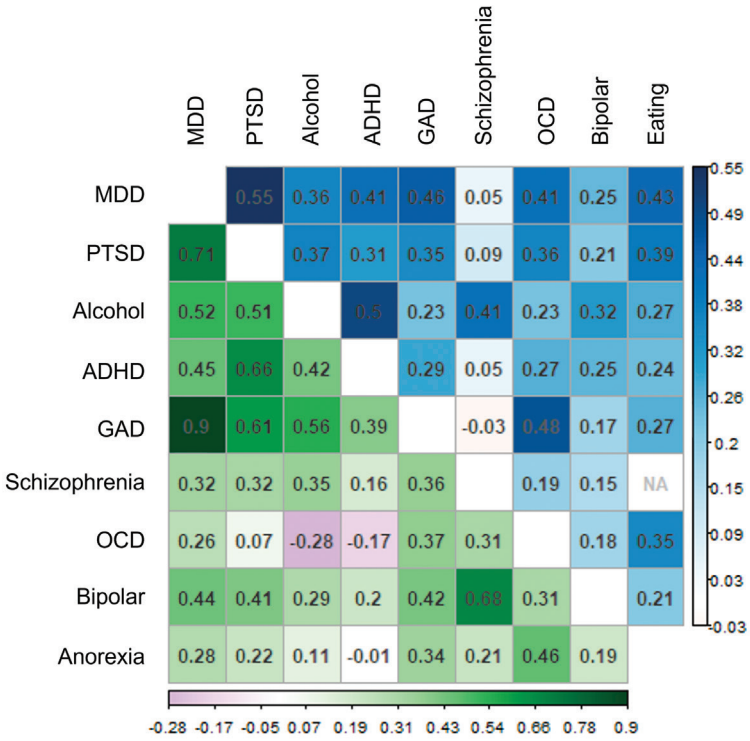


Figure 1 | Phenotypic and genetic correlations between diagnoses. Upper diagonal: polychoric correlations between diagnoses in the subset of 1.7 million siblings for whom education data was available in the Dutch population registry. NA: missing value due to low cell size. In this corner, anorexia nervosa was combined with bulimia nervosa into an eating disorder group due to the small number of individuals for either anorexia or bulimia diagnosis. Lower diagonal: genetic correlations between diagnoses estimated from summary statistics of the GWASs used in the MR analyses. Only disorders with data available in the Dutch population register and in GWASs are represented (see Supplementary Tables 6 & 25 and Supplementary Figure 2).

Figure 2 plots the prevalence (expressed in percentage) of each diagnosis given EA, split by sex (see also, Supplementary Table 5 & Supplementary Figure 2). Most diagnoses have a clear sex difference in prevalence, with higher prevalence for women, except for ADHD, Alcohol-related disorders, Schizophrenia. Most diagnoses show a decrease in prevalence with an

increase in education, from 11 years to 22 years of education. The group with 11 years of education stands out in two ways. First, they started in a pre-university track (selective track) in secondary school but dropped out without obtaining a diploma or re-orienting. Second, this group has the highest prevalence for having any diagnosis, including a particularly high prevalence of bipolar disorder and schizophrenia diagnoses. 0.75% of individuals with 11 years of education are diagnosed with bipolar disorder (below 0.4% in all other EA groups), and ~4% of men with 11 years of education are diagnosed with schizophrenia (below 2% in further EA groups). Speculatively, dropping-out without resuming formal education could be related to prodromal symptoms for these disorders.

Within-sibship analyses

As expected, simple logistic regressions of EA on diagnoses revealed a negative association between EA and being diagnosed, for all disorders except GAD (OR = 0.99, *SE* = 0.004) (Figure 3, Supplementary Tables 8-9). We then ran the within-sibship regressions: we regress diagnosis status on the average EA for all siblings in a family and the deviation of the sibling's EA from the family average. With these more robust within-sibship regressions, most associations were weaker (OR closer to 1), but still significant. On the other hand, for GAD, schizophrenia, bipolar disorder and anorexia, the within-sibship association were stronger (OR further from 1), the relationship between EA and diagnosis is stronger within than between families. Overall, odd ratios ranged from 0.95 to 0.81 per year of education, indicating a modest relation between EA and risk of diagnoses (Figure 3, Supplementary Figure 5, Supplementary Tables 10-11).

We run several post-hoc robustness analyses of the within-sibship regression. As for most diagnoses the prevalence differs by sex, we replicate our analysis in subsets of sibships with same-sex siblings only. For almost all diagnoses, the direction and magnitude of effects are the same. Bulimia is a notable exception: within-male sibship analysis suggests a positive relation with EA (OR = 1.25, *SE* = 0.12), but this relation is not significant (P = 0.02) (Supplementary Figure 4 and Supplementary Tables 12-13). We ran analyses excluding siblings with 11 years of education and with 2 years of education (which is an implausible outcome in the Dutch system). The results are qualitatively similar, but there are subtle changes in effect sizes when excluding siblings with only 2 years of education (Supplementary Figure 5, Supplementary Tables 12 & 14). Estimates of the relation between EA and schizophrenia diagnosis are the most sensitive to these exclusions.

We replicate this analysis using an alternative measure of mental care in the Dutch system: mental care expenditures, expressed in log(euro). Care expenditures include costs incurred within the specialized care, but also within the basic care (e.g. GP reporting mental health

care as the reason for visit). Mental care expenditures are also negatively associated with EA (β = -0.11, *SE* = 0.00). This association is reduced within-sibship (β = -0.08, *SE* = 0.00). Comparing same-sex siblings only, the within-sibship estimate is stronger in men (-0.1) than in women (-0.06). These effects correspond to 10% and 6% decrease in expenditure per year of education within-sibship (Supplementary Tables 16-17).

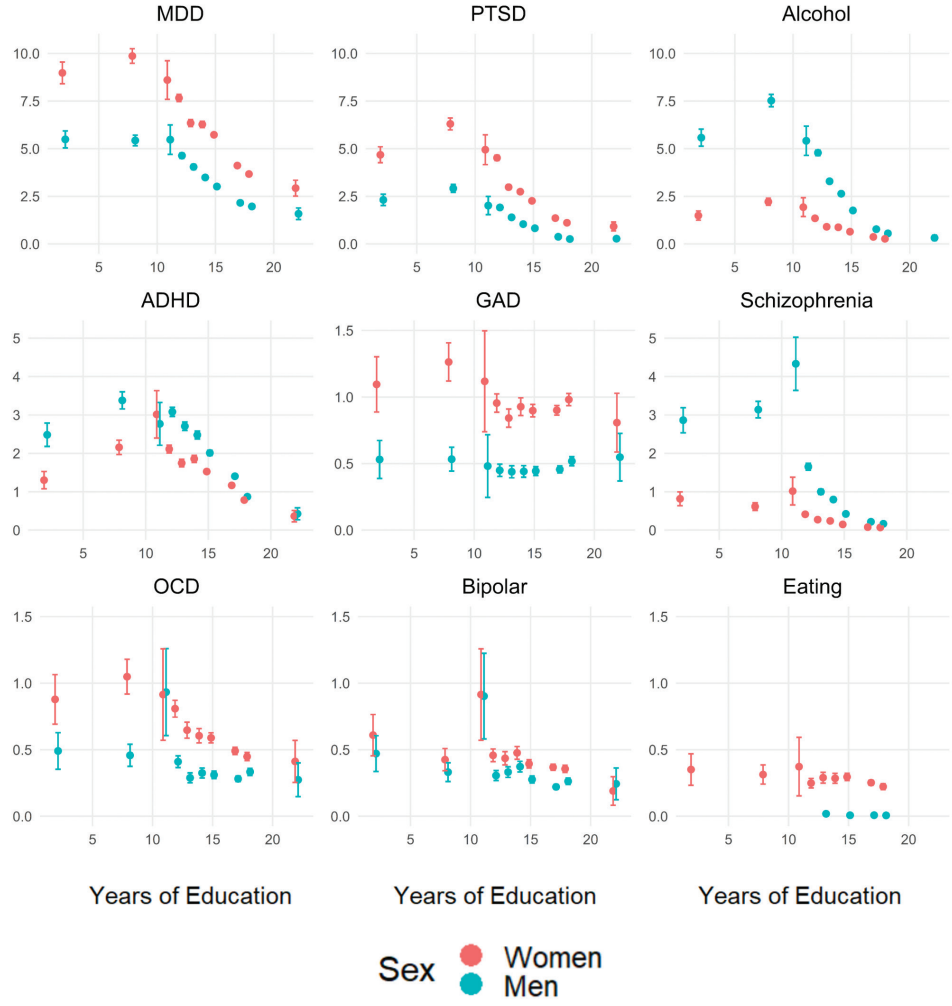


Figure 2 | Prevalence of diagnoses given someone's educational attainment and sex. Prevalence (expressed in percentage) of diagnoses in the sample of siblings for whom education data was available stratified by disorder (panels), education (x-axis) and sex (colour). Bars represent 95% confidence intervals. Note the scales of the y-axis are adapted depending on the diagnosis. In this figure, anorexia nervosa was combined with bulimia nervosa into an eating disorder group due to the low sample size of diagnosed individuals for some EA/sex strata. Only disorders with data available in the Dutch population register and in GWASs are represented (see Supplementary Table 5 and Supplementary Figure 3).

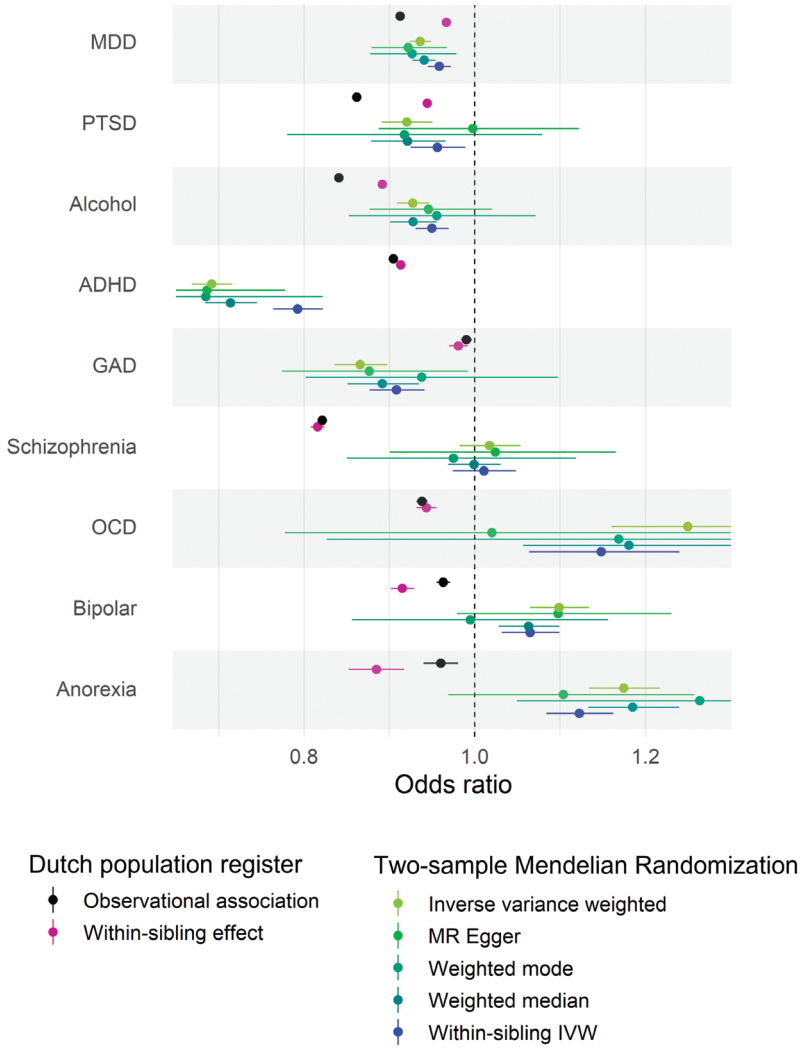


Figure 3 | Relation between EA and diagnoses as estimated with logistic regression, within-sibship regression or MR. Odds ratios per year of education as estimated with logistic regression (black) and within-sibship models (purple) in the Dutch population register and the two-sample MR analyses of EA on diagnoses (green/blue). Bars: 95% CIs. Only disorders for which both results are available are represented (see Supplementary Tables and Supplementary Figures 4 & 7).

Two-Sample Mendelian Randomization

MR estimates of EA on psychiatric diagnoses

With EA as exposure and relatively strong instruments (mean F-statistics > 50), MR IVW estimates are not always consistent with within-sibship estimates (Figure 3, Supplementary

Tables 20-21, Supplementary Figure 3). They suggest a comparable causal effect of EA on diagnosis for MDD, PTSD, and Alcohol dependence (IVW *OR* = 0.94, 0.92, and 0.93, $p < 0.004$). For ADHD and GAD, the risk-decreasing effect of EA estimated in MR is stronger than estimated in the within-sibship design (IVW OR = 0.69 and 0.87, $p < 0.004$). Importantly, MR estimates were yet not supportive of a protective effect of EA on four diagnoses. MR shows no effect of EA on schizophrenia (IVW *OR* = 1.02, $p = 0.33$), and a risk-increasing effect of EA on diagnosis for bipolar disorder, anorexia, and OCD (respectively IVW *OR* = 1.10, 1.17, and 1.25, $p < 0.004$), such that the liability to higher EA is causing a higher liability for these diagnoses. MR-Egger, weighted-mode and weighted-median estimates were mostly consistent in direction and effect sizes. Strick interpretation of evidence from pleiotropy-robust methods confirms a causal effect of EA on MDD and ADHD but not PTSD, GAD, or alcohol dependence. I2 were larger than 0.9 suggesting MR-Egger's estimates could be interpreted. There was evidence of heterogeneity for each EA-disorder pair indicating pleiotropy (Cochran's *Q* between 821 and 2101, *ps* < 0.004) but non-significant MR-Egger intercepts suggest the estimates are not biased by horizontal pleiotropy.

Exploratory sensitivity analyses based on SNP-effects on EA from a within-sibship GWAS yield lower estimates in the same direction for MDD, Alcohol dependence, ADHD and GAD (estimates became non-significant for PTSD and OCD) (Supplementary Table 21). The reduction in effect size is particularly large for ADHD (IVW *OR* = 0.69; within-sibship EA IVW *OR* = 0.81). However, the instrument was weak (mean F-statistics ~10.5), which can bias MR results towards the null.

MR estimates of psychiatric diagnoses on EA

When considering the reverse effect of mental disorder liabilities on EA, we relaxed the p-value threshold for instrument inclusion to 1e-5 for GWASs with low number of genome-wide significant hits (GAD, OCD, PTSD, alcohol dependence). The mean F-statistics for the instruments were modest (34.8 to 44), and modest-to-weak (21.5 to 22.8) when the *p*-value threshold was relaxed.

IVW estimates suggest a negative causal effect of the liability to disorder on education attainment (hence a bidirectional negative effect) for ADHD (IVW $\beta = -0.38$, $p < 0.004$), and at $p < 0.05$ (a liberal threshold given the number of MR tests) for MDD (IVW $\beta = -0.32$, $p = 0.010$), PTSD (IVW $\beta = -0.06$, $p = 0.026$), and GAD (IVW $\beta = -0.09$, $p = 0.016$) (Figure 4, Supplementary Table 22-23). We find no evidence of an effect of schizophrenia (IVW $\beta = 0.00$, $p = 0.94$), while higher liabilities for bipolar disorder and anorexia have a positive effect on EA (IVW $\beta = 0.17$ and 0.26, $p < 0.004$). The estimated effects of alcohol dependence (IVW $\beta = -0.05$, $p = 0.18$) and OCD liabilities (IVW $\beta = -0.00$, $p = 0.63$) on EA were not significant.

Again, MR-Egger, weighted-mode and weighted-median estimates were mostly consistent in direction and effect sizes. However, MR-Egger's estimates had low confidence and the weighted-mode estimate was <0 (non-significant) for bipolar disorder. There was evidence of heterogeneity of effects for every disorder-EA pair (Cochran's *Q* between 26.3 to 2602.1, *ps* < 0.004), but for OCD-EA (Cochran's *Q* = 29.6, *p* = 0.04). While the MR-Egger intercepts were never significant, the intercept for anorexia-EA was the highest (0.06, *SE* = 0.01), which might indicate the presence of horizontal pleiotropy. All I2 were larger than 0.9, however the small number of instruments for most diagnoses (e.g. anorexia N SNPS = 4) results in poor resolution to resolve directional pleiotropy. Exploratory analyses with SNP-effects based on within-sibship EA GWAS were consistent in direction and effect size with the previously described IVW estimates.

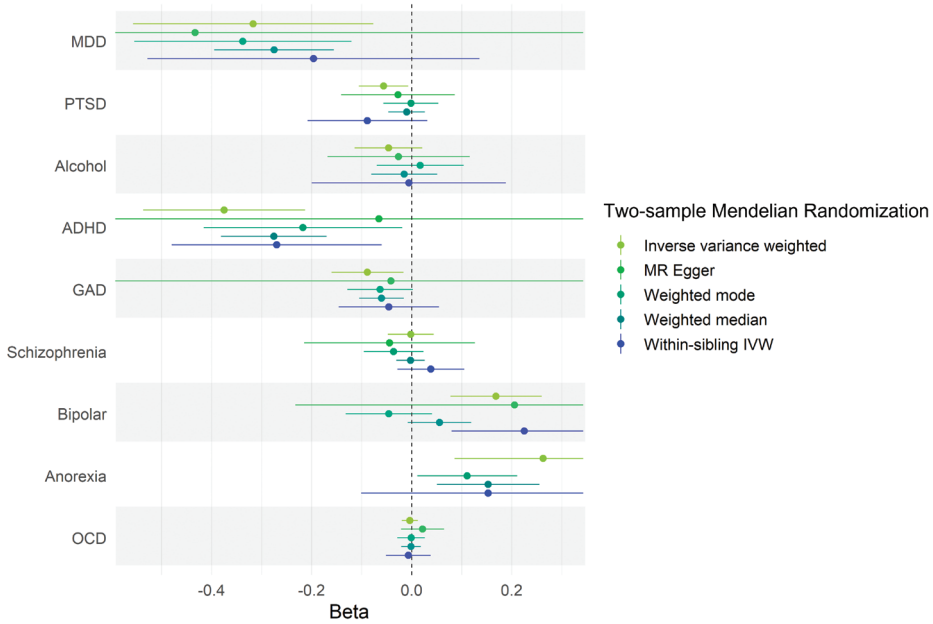


Figure 4 | MR analyses diagnosis as exposure and educational attainment as the outcome. Effect estimates of the two-sample MR analyses of EA on diagnoses. Bars: 95% CIs. Only disorders with data available in the Dutch population register and in GWASs are represented (see Supplementary Tables 22 & 23 and Supplementary Figure 4).

Education of healthy siblings in the Dutch population registry

Motivated by the discordance between the within-sibship and MR analyses for bipolar disorder, anorexia, and OCD, the first implying a protective effect of education and the second a risk, we investigate the hypothesis that the (genetic) liability for being diagnosed with some disorders is in fact associated with higher education, while the disorder itself (or

its prodromal manifestation) interferes with schooling. Under this model, bipolar, anorexic and OCD patients are expected to have a higher familial and genetic liability for EA than the general population. Therefore we expect the healthy siblings of these patients, but not of other patients, to be more highly educated than individuals in unaffected families. In the Dutch register, we do find that siblings of patients have less education than siblings in unaffected families for all disorders but bipolar disorder and anorexia (Figure 5, Supplementary Table 24). Unaffected siblings of patients with bipolar and anorexia have the same average years of education as unaffected families (average EA of unaffected families = 15.54; bipolar = 15.58, anorexia = 15.75, t-test p = 0.22 for bipolar, 0.006 for anorexia). Siblings of Bipolar II patients had slightly higher EA than individuals in unaffected families (15.68, $p < 0.004$) (Supplementary Figure 6).

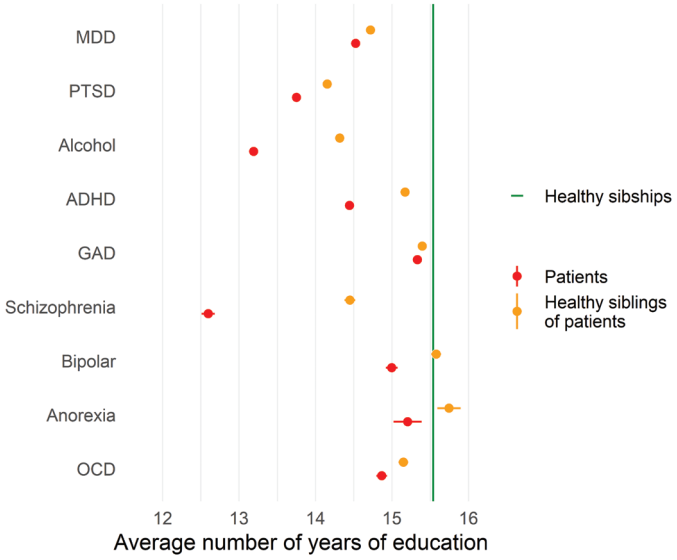


Figure 5 | Average number of years of education of patients, of healthy siblings of patients and unaffected sibships. The green line is the average number of years of education of unaffected sibships (15.54 years of education), in families in which none of the siblings was diagnosed with a mental disorder between 2011 and 2016. Red dots are the mean education of affected siblings per diagnosis. Orange dots are the mean education of siblings of an affected sibling that are themselves were never diagnosed in the 2011-2016 timeframe. Bars: 95% CIs. Only disorders with data available in the Dutch population register and in GWASs are represented (see Supplementary Table 24 and Supplementary Figure 8).

DISCUSSION

Triangulating across two designs, we find consistent protective effects where higher EA reduces the risk of MDD, alcohol dependence, GAD, ADHD, and PTSD diagnoses, though for some traits, there is also evidence for reverse causation. Critically, for MDD, alcohol

dependence and PTSD, the implied causal effects in MR and within-sibship designs are substantially lower than the observational association. For other diagnoses, notably bipolar disorder, anorexia and OCD, the two designs yield inconsistent results. These are noteworthy and gave rise to additional analyses. Our study reveals that statistically convincing findings based on either design would have provided an incomplete or misleading understanding of the relationship between education and mental health.

First, we discuss the findings for some common diagnoses, highlighting representative results. Similar interpretation and integration of the broader literature could be done for the other diagnoses, but we defer to scientists with specific expertise on these disorders. We close with limitations and conclusions

Depression

Both the within-sibship and the MR approach suggest around 6% lower odds of depression diagnosis per additional year of education, while there is also suggestive evidence for reverse causation (depression diagnosis or liability hampers EA). The implied causal effect of education on the risk of depression diagnosis is lower that the observational relationship. Our results are consistent with work that relies on other quasi-experimental methods, for example exploiting compulsory school-law reforms lengthening minimum EA[244]. One study found a ±7% reduction in the probability of meeting diagnostic criteria for depression (N = 3,704)[245]. Another study (N = 21,085) found no significant effect of EA on depression, though their negative sign is consistent with our findings[246,247]. Multiple previous (MR) studies based on UK Biobank data, which is an older (age 40-70) and more highly educated cohort, find no effect[246,248]. The inconsistency with the previous literature might be due to their samples of healthy volunteers, while our within-sibship sample is population-based. An additional year of education might especially benefit the mental health of those at the lower end of the education distribution (a group under-represented in volunteer samples)[244]. Also, the protective effect of EA on mental health might wane with age[231].

We focussed on individual's final obtained degree. However, EA isn't the only aspect of education that relates to mental health. For example, a German reform reduced the academic high-school track by one year while keeping the curriculum constant. This intensified schooling increased depressive symptoms[249]. Others have also shown that kids being relatively young for their school grade are more likely to be diagnosed with depression[250], or that later separation of students into different tracks might increase depression in women[251]. While we find that higher EA plausibly reduces the risk of many diagnoses in adulthood, any reform aimed at increasing EA should keep in mind that other aspects of education might have opposite effects.

ADHD

For ADHD reciprocal processes are plausible. The formal diagnosis of ADHD alludes to problems that interfere with school work. This creates an obvious dependence between ADHD and poor performance in school (socially or academically). Other two-sample MR studies also reported the bidirectional effect of EA and ADHD. Additionally, they suggest this effect is independent of IQ[239,252]. Interestingly, our MR estimate of the effect of EA on ADHD is lower when using SNP-effects based on within-sibship EA GWAS, suggesting that assortative mating, population stratification, and/or gene-environment correlation might bias MR estimates. These biases could explain the difference in estimates between MR and the within-sibship analysis. On the other hand, the difference between the observational and within-sibship association registry is small, suggesting familial factors do not explain the association between EA and ADHD diagnosis.

ADHD is part of the "disorders usually first diagnosed in infancy" in the DSM-IV, hence usually diagnosed at school age, which seem inconsistent with a causal effect of EA on ADHD diagnoses. Our data concerns ADHD diagnoses in adulthood. As we do not have data on age at first diagnosis, we cannot distinguish between a recurring diagnosis and first diagnosis. However, there are plausible mechanisms through which EA might continue to affect the likelihood of diagnosis after one's educational trajectory. For example, higher-educated individuals might better cope with their symptoms and might have more power to select suitable environments (e.g. selection of jobs with more flexibility[253]).

Schizophrenia

The within-sibship design estimates 18% lower odd of schizophrenia per additional year of education, while MR suggests no causal relationship. The within-sibship association might reflect the interfering effect of prodromal or early symptoms of schizophrenia on education. This would be consistent with our observation that 4% of Dutch men who drop out of the pre-university high school track, and do not re-enter education, are diagnosed with schizophrenia. However, we do not observe a causal effect of schizophrenia on EA in our MR analysis. Note that the interpretation of MR is difficult when the exposure is binary and rare. The MR estimates should be viewed as reflecting the effect of the liability for being diagnosed, not an effect of the diagnosis itself. It is easy to imagine how core symptoms of schizophrenia, like delusions, hallucinations, and disorganized thoughts, interfere with education, but these symptoms are not gradually experienced over the entire liability spectrum. If we do accept the MR findings, the discrepancy between the within-sibship and the MR results could be alternatively explained by factors not shared between siblings

that reduce EA and increase the risk of schizophrenia diagnosis (e.g. risk exposures such as trauma). Contrasting our findings, previous MR estimates using earlier schizophrenia GWASs hint at an increased risk for schizophrenia per additional year of education[238,240], as we found for anorexia, OCD, and bipolar disorder.

Bipolar disorder

Bipolar disorder, anorexia nervosa, and OCD have the most striking pattern of results. While the within-sibship design suggests that each additional year of education reduces the likelihood of being diagnosed by 6 to 12%, the MR estimates suggest that EA increases the likelihood. We hypothesised that these diagnoses follow a model where the (genetic) liability for being diagnosed is associated with higher education, while the disorder itself (or its prodromal manifestation) interferes with schooling. Supporting this, we found (1) a high prevalence of bipolar disorder among pupils dropping out of the pre-university high-school track and (2) healthy siblings of bipolar and anorexic patients have similar or even higher EA than the average unaffected sibships. Other Dutch studies comparing siblings of patients confirm this observation for bipolar disorder compared to schizophrenia or MDD[254,255]. However, this was not replicated in Denmark[256]. Bipolar disorder is associated with traits that are valuable in school settings like more creativity[257,258] and higher childhood IQ[259]. A Swedish population-cohort study reports that individuals with excellent high-school performance have a fourfold increased risk of bipolar disorder over those with average grades[260]. A true positive relation between factors that increase both success in education and risk of bipolar diagnosis seems supported, although the mechanism remains unclear. They could share biological mechanisms or psychological traits like creativity. Likewise, higher-educated individuals, or individuals coming from higher-educated families, could be more likely to be diagnosed with these disorders, due to more proactive help seeking, better access to care that facilitates these diagnoses, or preferential diagnosis by practitioners. Finally, biased selection into GWASs could result in the selection of bipolar disorder or anorexia cases with higher-than-average education. Further studies are essential to understanding the mechanisms of the apparently contradictory association between these disorders and EA.

Limitations

Our conclusions rely on within-sibship and MR estimates, and on whether we can interpret contradictory findings. We assume these contradictions are mainly due to true differences in underlying phenomena. However, other statistical, measurement-, and sampling-related factors could play a role.

The two methods rely on data obtained in different samples drawn from different

populations. Our within-sibship analysis was based on registry data, containing most of the Dutch population born between 1985-1965, so highly representative for the Netherlands. In contrast, the MR is based on GWASs of international samples who are volunteers of European-ancestry. Both higher risk for mental illness and lower EA are known to increase non-participation and participant attrition[261–264], a selection bias that could induce a collider bias[265]. In the within-sibship analysis, we rely on diagnoses by professionals in specialized care, therefore dependent on the current referral policies of the Dutch healthcare system. All our GWASs rely on diagnoses, but they differ on the report (self-report, medical files, etc.) and on the timeframe (most look at lifetime diagnoses, but the age range of participants vary widely). In both types of data, who seeks help and who is diagnosed may depend on class, ethnicity and context. Our conclusions are phrased in terms of diagnoses; potential mechanisms discussed are speculative. Any conclusions beyond diagnoses run into the imperfect relation between people's symptoms and people's DSM-IV diagnosis.

Regarding two-sample MR, we use MR analyses robust to weak-instrument and pleiotropy, but the additional statistical power required make the estimates more uncertain. To control for potential biases due to demographic and dynastic effects, we used summary statistics from a within-sibship GWAS. However, EA is the only trait with a within-sibship GWAS suitable for MR, leading to an unbalanced control of these biases in our MR. Notably, dynastic effects could lead to false positive bidirectional effects[266]. Similar well-understood caveats apply to sibling designs but most additional criticisms of this design[221,222] pertain to lack of insight on the direction of effect and the possible underestimation of effect sizes.

Finally, our findings pertain to the effects of education on mental health within the current confines of social life and social policy. Our findings do not identify an immutable or permanent cause of differences in mental health. Policy changes outside education (e.g. minimum wage, affordable living, quality housing) may improve people's mental health and could as effectively close the health gap between educational groups.

Conclusion

The aforementioned caveats limit the certainty we should ascribe to specific causal claims. Nevertheless, we established consistent potential causal effects of EA on the risk of being diagnosed with MDD, ADHD, PTSD, alcohol dependence, and GAD. These potentially causal effects are smaller than the observational associations (but for GAD). For diagnoses like bipolar disorder and anorexia, our results suggest a positive relationship between EA and the diagnosis liability, yet a negative relationship between EA and the diagnosis itself. These patterns deserve further study and would have been missed when applying a causal-inference technique in isolation.

METHODS

This study was pre-registered at: https://osf.io/vmpfg/?view_only=b17c64b5600c4d32902e55ea26d63f37. Deviations to the preregistration are detailed in Supplementary Note. All code associated with the analyses is available on GitHub at https://github.com/PerlineDemange/CBS-MR. We follow the STROBE[267] and STROBE-MR[268] reporting guidelines (Supplementary Tables 26 & 27). This research was reviewed and approved by the Scientific and Ethical Review Board (VCWE) of the Faculty of Behaviour & Movement Sciences, VU University Amsterdam; application number VCWE-2020-054.

Within-sibship design

Data source

We analyse restricted access microdata from Statistics Netherlands (CBS). Under strict conditions, these microdata are accessible for statistical and scientific research. For further information on remote access procedures: microdata@cbs.nl.

Study population

We included individuals born between 1965 and 1985 (N = 6,539,767), such that they are between 26 and 46 years old when the first year of diagnostic data is available. From these we select siblings (sharing the same legal mother and father), where more than one sibling has educational data available. We retain a final sample of N = 1,743,032 individuals nested within 766,514 families. For a comprehensive overview of the selection procedure, see Supplementary Note and Supplementary Figure 1.

Educational attainment

Educational attainment data is based on various registers and surveys and has a high coverage (more than 11 million people). Based on the final degree obtained we inferred the number of years of full-time education of the individual. The transformation of the 17 diploma categories to years of education is available in Supplementary Table 1. For a comprehensive overview of the variable definition, see Supplementary Note.

Mental health outcomes

The **Dutch mental health care** system distinguishes two systems of care. Here we rely on diagnostic data for specialized/second-line care. Specialized mental care is intended for patients with severe or complex diagnoses which require the attention of a psychiatrist or clinical psychologist. **Psychiatric diagnoses** are obtained from the care trajectory of patients getting specialized mental care. Diagnoses are classified based on the Diagnostic

and Statistical Manual of Mental Disorders 4th edition (DSM IV). We consider an individual as affected if they were diagnosed with any of the disorders listed in Table 1, in at least one year during the 2011 to 2016 period.

Mental health care expenditures[269] are assessed in two ways: expenditures from the first line/basic care and expenditures from the 2nd line/specialized care. We summed basic and specialized care for each year. Due to the steeply skewed distribution of incurred mental health expenditures (Supplementary Table 16, Supplementary Figure 6), we averaged the expenditures across 2009-2018 for each individual and log-transformed the personal average.

Statistical analyses

All diagnoses analyses were done in the sibling sample (4), for each mental health diagnosis separately. We estimate polychoric correlations between all diagnoses. We ran an observational analysis: a logistic regression with EA as a predictor and the psychiatric diagnosis as outcome, ignoring family structure. We then ran within-sibship logistic regression analysis. We regress diagnosis status on the average EA for all siblings in a family and the individual deviation of the sibling's EA from their family average. The effect of the average EA (between-sibship effect) represents the expected change in the outcome being diagnosed given a one-unit (the unit being scaled as a year of education) change in the sibling average, while the effect of the deviation of the sibling to their family average (within-sibship effect) represents the effect of EA when keeping the factors common to the family constant. We do not correct for family clustering in these analyses. Part of the reason is the computational limitation of logistic regressions with random family effects in large sample sizes and low prevalence. If the dependence between errors for relatives is very strong we approach an upper bound where the number of families, rather than individuals, equals the number of independent observations. As most families in our data consist of 2 siblings, we can estimate the upper bound for the SEs to be sqrt(2)*uncorrected SEs. As our uncorrected SEs are very small, this correction would not change our results. Additionally, we focus on interpreting the effect sizes, p-value being of little interest at this sample size. For the analysis of mental health care expenditures, we fit a linear model instead of a logistic model as the outcome is continuous, and also report results from linear models with random family effects. In all analyses, we included sex, birth year and birth order as covariates.

As sensitivity analyses, we separately considered men and women (from same-sex sibships) and we omitted specific education groups that are rare or implausible given the Dutch educational system (11 and 2 years of education). We investigated mean differences in EA between patients, siblings of patients, and families that are entirely unaffected. We performed t-test to compare the mean EA of these groups for each disorder.

Two-sample Mendelian Randomization

We follow recommendations by Burgess et al[270].

Summary Statistics

We relied on summarized statistics from a large well powered 2018 GWAS of EA[21] (EA3). We reproduced this GWAS by meta-analysing published summary statistics with summary statistics obtained from 23anMe, Inc, as done in [152]. For additional sensitivity analyses, we relied on summary data from the within-sibship GWAS of EA[229]. The within-sibship GWAS is significantly smaller, but because the SNP-effects are estimated within-family they are unbiased by potential effects of assortative mating, population stratification or intergenerational genetic effects[271,272].

For each psychiatric disorder, we selected GWAS summary statistics preferentially selecting the most recent or largest GWAS available. A full list of GWASs and a description of the summary statistics are available in Supplementary Table 18. We assessed potential sample overlap between EA and psychiatric disorders GWASs using LD-score cross-trait intercept as proxy[273] (Supplementary Table 19 & Supplementary Note).

Choice of the genetic variants

For data cleaning and analyses we used TwoSampleMR[274] in R.4.1.0[275]. When analysing EA as the exposure, we first excluded genetic variants not present in the outcome summary data. For EA3 summary statistics, we selected genetic variants associated with EA at $p <$ 5e-8. For within-sibship EA summary statistics, we selected significant independent loci identified by the EA3 GWAS, which were also associated at p-value < 0.05 in the within-sibship EA GWAS. We then clumped to select statistically independent variants (kb = 1000, r2 = 0.001). When analysing psychiatric disorders as the exposure, if clumping genetic variants associated at $p <$ 5e-8 led to the selection of less than 5 genetics variants, we further selected genetic variants associated with the exposure at $p <$ 1e-5 (this occurred for autism, GAD, bipolar-II disorder, OCD, PTSD, and alcohol dependence). Variants were then harmonized between exposure and outcome summary statistics, ensuring the SNP effect relates to the same reference allele. Ambiguous and palindromic variants with MAF > 0.42 were excluded.

We scaled the effect sizes from the two EA GWASs so that the SNP-effects reflect change in term of years of education. For this, we estimated the weighted average standard deviation of the education phenotype in the cohorts included in the EA3 study (SD = 3.9 years) and multiplicated both SNP-effects and their SEs by this number[239,276]. Estimates for the psychiatric disorders SNPs were converted to log(OR) if reported in OR. MR estimates were

transformed back to the OR scale where needed.

Analyses

We ran two sets of MR analyses: EA as exposure and mental disorder as an outcome, and mental disorder as exposure and EA as an outcome. We ran an inverse-variant weighted (IVW) mendelian randomization[225]. We judged the significance of the p-value following a Bonferroni correction: significance threshold 0.05/12 traits = 0.004. To test for the robustness of the IVW findings against potential violation of the MR assumptions, we ran MR-Egger[226], weighted-mode[228], and weighted-median[227] analysis. We reported the Cochran Q's-statistic SNP effect heterogeneity and the F-statistics assessing potential weak instruments bias[277,278]. Additionally, we reported the I2 statistic[279], which gives an indication of the violation of the NO Measurement Error (NOME) assumption, on which MR-Egger relies. We also report LD-score based genetic correlations between all GWASs, computed with Genomic-SEM.

CODE AND DATA AVAILABILITY

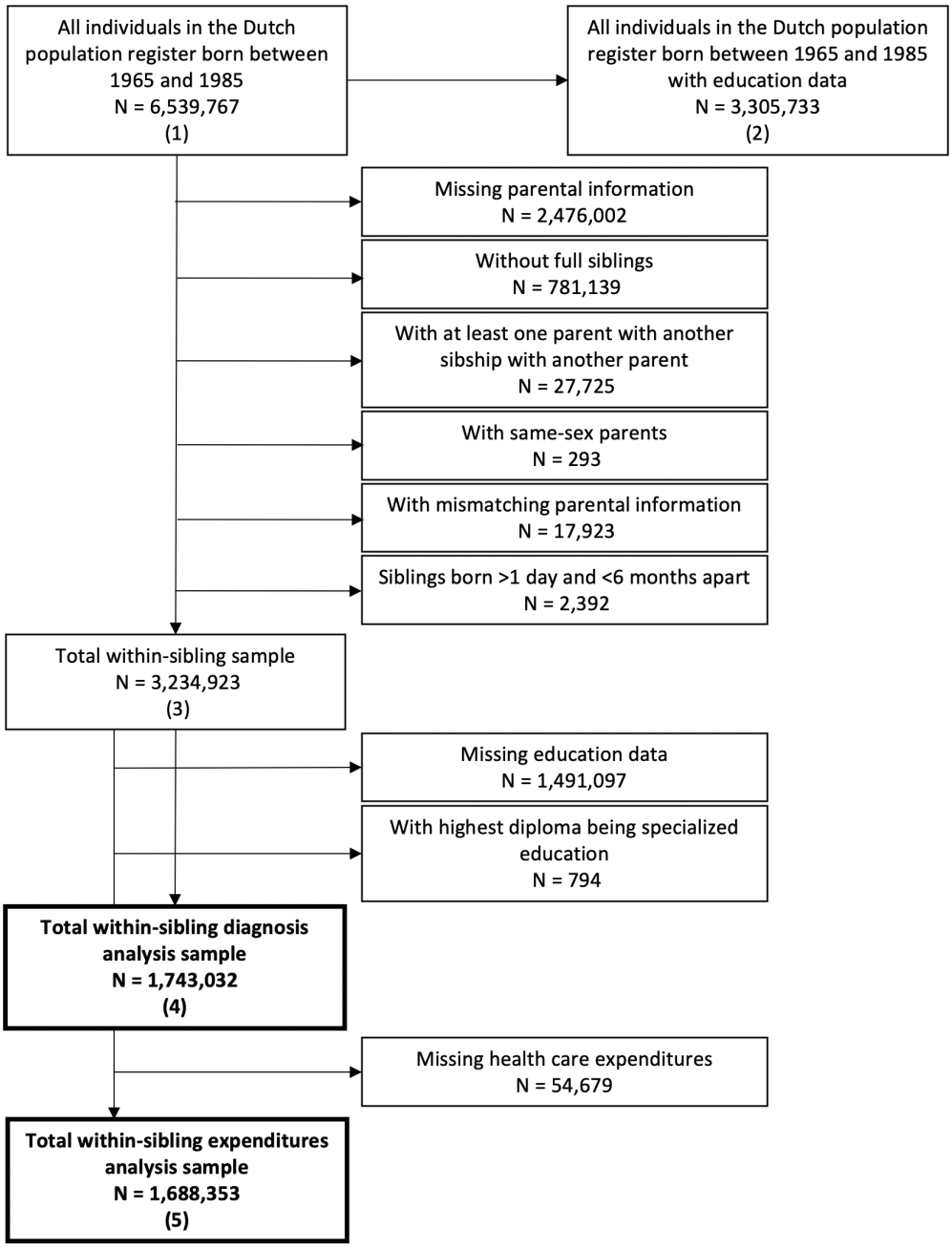
All code associated with the analyses is available on GitHub at https://github.com/PerlineDemange/CBS-MR.

We analyse restricted access microdata from Statistics Netherlands (CBS). Under strict conditions, these microdata are accessible for statistical and scientific research. For further information on remote access procedures: microdata@cbs.nl. For GWAS summary statistics availability, see original publications. For 23andMe, Inc. dataset access, see https://research.23andme.com/dataset-access.

SELECTED SUPPLEMENTARY INFORMATION

Full Supplementary Information and Supplementary Tables can be downloaded at:

https://www.medrxiv.org/content/10.1101/2023.01.26.23285029v1.supplementary-material



Supplementary Figure 1 | Flowchart of study sample

CHAPTER 5

NO EFFECT OF PARENTAL MENTAL HEALTH ON CHILDREN'S ACADEMIC ACHIEVEMENT: A WITHIN-FAMILY STUDY

In preparation as: **Demange, P.A**., Nivard, M.G., Torvik, F., Eilertsen, E.M., Cheesman. R., Nordmo, M., Lyngstad, T., Ystrøm, E.*, van Bergen, E.* No effect of parental mental health on children's academic achievement: a within-family study

Supplementary materials accessible at:

https://osf.io/sqz3n/?view_only=2b34df0ee22c484d85682ad23fa4db52

ABSTRACT

Children of parents with psychopathology generally do less well in school than their peers. Parental symptoms might however not be responsible for their lower achievement, as other familial factors might be at play. To examine the role of parental symptoms, we analyse data from up to 9,000 families of the Norwegian Mother, Father, and Child Study (MoBa). Parents filled out surveys on their symptoms of anxiety, depression, eating disorders, ADHD, and alcohol use disorder. Children in 5th Grade (aged 10) participated in nationally-standardised tests of mathematics, reading comprehension, and English (as an additional language). Comparing families whose parents are siblings controls for unmeasured factors shared among adult siblings (e.g. genetics and socioeconomic status) that confound the relationship between parental mental health and children's academic achievement. Simple regressions, not controlling for familial confounding, showed that children tend to score slightly lower on maths and reading if their parents had more symptoms of anxiety, depression, or eating disorder ($-.060 \leq \beta s \leq -.014$). However, these associations were attenuated and no longer significant within families. While parental psychopathology symptoms correlate weakly with children's academic achievement, our findings suggest that these correlations are due to familial confounding. That is, our findings suggest no causal effect, or at least no substantial effect, of parental mental health on children's achievement. Our study highlights the value of within-family designs to understand the causes and consequences of psychopathology.

INTRODUCTION

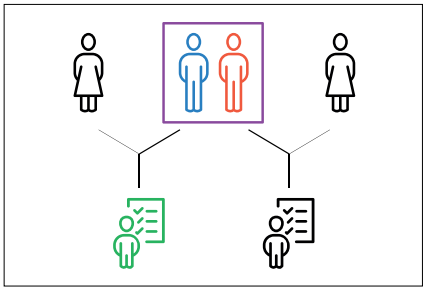
A significant proportion of children live with parents suffering from mental disorders: recent assessments in western countries suggest 10 to 25% of children have an affected parent[280–283]. Children of parents with mental disorders are a particularly vulnerable population. Parental psychopathology is associated with childhood adversity[283] and poorer child health, behavioural and academic outcomes[284–286].

Overall, previous studies suggest that poor parental mental health is associated with poor educational outcomes in children. However, evidence for a negative association is highly inconsistent across studies and might depend on the specific parental disorder, its severity, time at measurement, children's outcomes, parent gender, study population, etc. Internalizing disorders are among the most frequently studied parental risk factors in this regard, and studies report varying degrees of null[287–289] to negative[287,289–294] associations with children's academic achievement. Having a parent with schizophrenia also appears to negatively affect children's academic performance, while having a bipolar parent does not[292,295]. The effect of parental eating disorders has yielded divergent results[296]. The association with alcohol and substance use is also unclear and might depend on timing of problematic usage and quantity[297,298].

All these studies report observational parent-child associations, adjusted for a limited number of measured covariates. While identifying specific negative associations between parental mental health and academic achievement allows to identify at-risk groups, identifying causal effects is the only thing that could inform on mechanisms at play in the parent-child transmission, and on potential effects of interventions on parental mental health. Observational parent-child associations are likely to be confounded by familial factors, such as social-economic position and genetic influences[294,299]. Such familial factors can be controlled for in within-family designs: designs that employ family structure in pedigree data to reduce the impact of genetic and environmental confounding factors that are shared within a family. These design have the additional advantage to not require the identification of these factors, and therefore do not rely on the imperfect measurement of these when adjusting estimates. To our knowledge, only two studies use such designs to study the effect of parental mental health on children's education, looking at maternal smoking[300] and parental schizophrenia[301]. Both find no evidence for a causal effect on children's school performance.

Here, in a design which we refer to as children-of-siblings design (Figure 1), we use family structure to investigate effects within-family: we test whether the sibling with worse mental health is also more likely to have a child who has lower academic achievement. This design

allows us to control for unmeasured factors shared between siblings that might relate to their mental health and their children's education, such as a common environmental/social influence, social economic position or shared genetic influences. It therefore allows us to get an estimate controlled for various unmeasured sources of confounding which would persist in an observational association. While it is by itself not sufficient to conclude causality (confounders not shared between siblings might still confound association estimates), comparing (children-of-)siblings informs us on the presence or absence of causality more precisely than observational studies. Additionally a reduction in the within-sibling estimate compared to the observational estimate indicates that factors shared to the family have a role in the association between parental mental health and child achievement.



Is the sibling with worse mental health more likely to have a child who has lower academic achievement?

$$\begin{aligned} &Standardized\ test\ score_{ij} \\ &= \beta_0 \\ &+ \beta_{within}\left(Symptom\ score_{ij} - \overline{Symptom\ score_j}\right) \\ &+ \beta_{between}\overline{Symptom\ score_j} \\ &+ Z_{ij} + u_j \end{aligned}$$

Figure 1 | Visualization of the children-of-siblings design.

METHODS

The study was preregistered, the registration is available on the Open Science Framework website: https://osf.io/sqz3n/?view_only=2b34df0ee22c484d85682ad23fa4db52. Deviations from the pre-registration are stated in Supplementary Note.

Sample

We use data from the Norwegian Mother Father and Child Study (MoBa), release 12. MoBa's mothers were recruited all over Norway from 1999-2008 during their pregnancy. 41% of the contacted pregnant women consented to join the study and partners were asked to join. The cohort follows 112,645 pregnancies, 95,135 mothers and 75,111 fathers, including extended families.

We estimate associations in the full sample of the Norwegian Mother Father and Child Study (MoBa) with available measures, and within-family associations in a subset thereof (referred to as the children-of-siblings sample). We used linkage to the Norwegian central population register for basic demographic data and kinship links.

Within the Norwegian central population register, we identified all children sharing the same legal mother and father, excluding children with same-sex parents. We then identified children whose mothers are sisters (i.e. these children are maternal cousins, they share maternal grandparents), whose fathers are brothers, and whose parents are siblings of the opposite sex. Linking this information on cousins and parental siblings, we selected in MoBa parents for whom at least one full-sibling (including twin) also participated in MoBa. These parental siblings and their children were included in the analysis if both siblings reported mental health and at least one of their children has data available for the national academic test in 5th grade. We only included one child per parent. Some children were part of multiple extended families: once via their mother, once via their father. We excluded one of these familial links at random to avoid duplicated children. Our main analyses use the full sample of cousins in the child generation based on maternal, paternal and cross-sex parent siblings to maximize the sample size. Sample sizes are presented in Table 1. Pre-registered analyses were focusing on same-sex parent siblings, and full results are available in Supplementary Tables 7-9.

Measures

Academic achievement

The child academic outcome variables are individual scores on national standardized tests in 5th grade. Pupils take national standardized tests in 5th, 8th and 9th grade in mathematics, reading comprehension (in their native language: Norwegian) and English as an additional language. We focus on 5th grade (~10 years-old) as it is before the typical age of onset of many psychopathologies and is the academic achievement with the largest available sample size linked to MoBa. The test scores are obtained from linkage with the National Educational database (NuDB). We excluded pupils who took the tests in Sami language and who took the tests several times in the same year. For pupils who took the tests two years in a row, we only considered the results obtained the first year. We then standardized the scores within each test and year combination, in the overall NuBD population, so that our outcome measures how the child places themselves in their own cohort on the specific test[171].

Parental Mental Health

Extraction of the mental health results in MoBa was performed with the R package Phenotools[302] version 0.2.4. Parental symptoms scores were created following standard practice as implemented in MoBa and documentation of the questionnaires.

For all scores based on validated scales, we set a completion threshold of 0.75: 75% of the items had to be reported for the total score of the participant to be considered. Lower

completion was considered a missing score.

As a rule of thumb, we selected available parental measures assessed at 15 weeks of pregnancy (first MoBa survey and only survey sent simultaneously to both parents) and the closest in time to the child passing the national test (age 10). In mothers, anxious, depressive and eating disorders symptoms were assessed during pregnancy and at age 8, alcohol problematic use symptoms were assessed at age 8 and ADHD symptoms were assessed at 36 months. Fathers were surveyed only during pregnancy and in 2015. Anxious and depressive symptoms were assessed during pregnancy and in 2015, ADHD symptoms were assessed during pregnancy and alcohol problematic use symptoms in 2015.

Note that the 2015 questionnaire was sent regardless of the children's age, so this measure might be taken after the child's national test. For analyses with these measures, we only included the subset of fathers whose children were younger than 10 in 2015.

Anxiety and depression symptoms were measured using subsets of the Hopkins Symptoms Checklist (SCL)[303], we use SCL-5[304] at 15 weeks of pregnancy and SCL-8 at age 8 and 2015. ADHD symptoms were assessed with the Adult ADHD Self-Report Scale (ASRS Screener)[305,306]. Problematic alcohol use symptoms were measured with 7 items of the Alcohol Use Disorders Identification Test (AUDIT)[307,308] which relate to problematic use (AUDIT-P). Eating disorder symptoms according to the DSM-IV[309] diagnostic criteria for anorexia, bulimia, and eating disorder not otherwise specified (i.e. purging disorder and binge eating disorder) were only assessed in mothers[310]. With low prevalence of eating disorders and high level of comorbidity, we computed a continuous eating disorder risk score using item response theory on the DSM-IV eating disorder symptoms. In Stata 17, we computed a nominal response model for each time point, we report parameters estimates in Supplementary Tables 10 & 11 and test characteristics curve and test information functions in Supplementary Figure 1.

Statistical analyses

For each measure of mental health in parents, we run a series of models to test the association with childhood educational test scores. Each model aims to further control for measured and unmeasured confounding, through sibling comparison in the parental generation. We standardized the mental health scores with a variance of 1 and mean of zero before each model. For each model, we adjust our *p*-values for multiple testing, using a false discovery rate correction.

We first ran a mixed-effects regression to estimate the minimally adjusted association between the parental mental health traits and the child's academic achievement. We ran this

model (1) in the full MoBa parent-child couples with available measures and in the children-of-siblings subset.

$$EA_{ij} = \beta_0 + \beta_1 MH_{ij} + Z_{ij} + u_j \quad (1)$$

EA_{ij} is the standardized test score of the child i. MH_{ij} is the mental health score of the parent of the child i. Z_{ij} are covariates: sex and year of birth of the child, years of birth of the mother and of the father, sex of the parent. u_j is the random effect for the family j (family including the parental siblings and their children).

In the children-of-siblings sample, we estimated the intra-class correlation of mental health in parental siblings, as an indication of the proportion of variation in mental health between sibling pairs.

To estimate the association between the parental mental health score and the child's school performance adjusted for unmeasured shared familial factors, we then ran the children-of-sibling design: a regression including within and between-family effects of parental mental health on child's academic achievement, effectively using the siblings of parents as de facto matched controls.

$$EA_{ij} = \beta_0 + \beta_{within}\left(MH_{ij} - \overline{MH_j}\right) + \beta_{between}\overline{MH_j} + Z_{ij} + u_j \quad (2)$$

With $\overline{MHj}$ being the average mental health score of the parent and their siblings (average score for the family j).

The within-sibling effect represents the expected change in educational performance given one unit change in the difference between the parent's mental health and the average mental health of the parents' sibling pair. It represents the effect of parental mental health on the child's test score while keeping factors common to the extended family constant.

We also investigated the attenuation of the within-sibling effect when taking into account parental education, see Supplementary Note.

RESULTS

Descriptive statistics

Descriptive statistics for parental symptoms and child academic achievement are presented in Table 1 and Supplementary Table 1. To maximize sample size, we always use the full sample which differs depending on the missingness of these measures. Analytic sample sizes vary from 35,210 to 134,174 in the overall MoBa and from 2223 to 17,291 for the children-of-siblings subsamples. The average standardized test scores of MoBa children are above zero (0.11 to 0.34), indicating that children participating in MoBa have on average higher academic achievement than the Norwegian population (Supplementary Table 1). This is in particular for Maths and Reading scores, with average scores of 0.26 and 0.27 respectively, and even more pronounced in the children-of-sibling samples with average

scores of 0.33 and 0.34. Parents in the children-of-siblings samples had fewer symptoms and less variance in symptoms than the overall MoBa sample (Table 1). The intraclass correlation ranges between 0.09 and 0.16, indicating variation in symptoms load between siblings. For symptoms with measures at two time points, the intra-class correlation is lower when the child is 8 years old than during pregnancy (0.13 vs 0.10 for anxious and depressive symptoms and 0.16 vs 0.10 for eating disorders symptoms).

	Anxiety - Depression in pregnancy		Anxiety - Depression ~ age 8		ADHD ~ < 3yo		Alcohol Problematic Use ~ age 8		Eating disorder in pregnancy		Eating disorder ~ age 8	
Subsample	MoBa	CoS	MoBa	CoS	MoBa	CoS	MoBa	CoS	MoBa	CoS	MoBa	CoS
Scale	**SCL-5**		**SCL-8**		**ASRS**		**AUDIT-P**		**IRT**		**IRT**	
Sample N	134,174	17,291	61,852	4248	73,616	6835	61,473	4214	47,805	7850	35,210	2223
Mean	1.02	0.92	2.14	1.98	7.23	7.16	0.48	0.39	0.05	0.02	0.00	-0.03
SD	1.79	1.66	3.00	2.79	3.43	3.37	1.31	1.13	0.84	0.84	0.60	0.54
Skew	2.81	2.98	2.44	2.47	0.24	0.21	4.58	4.91	0.81	0.84	0.17	1.64
% girls	49	49	50	49	49	49	50	49	49	49	50	49
% mothers	55	57	56	67	61	65	57	67	100	100	100	100
Number of families	–	8427	–	2093	–	3361	–	2078	–	3855	–	1097
Intra-class correlation	–	0.13	–	0.09	–	0.10	–	0.12	–	0.16	–	0.10

Table 1 | Descriptive statistics of parental mental health at different time points in children's life. Average over the analytic samples (used in Figure 2). Statistics per sample are in Supplementary Tables 1 & 4. Subsample "MoBa" refers to all parent-child pair with non-missingness in MoBa, and "CoS" refers to the children-of-siblings subsample of MoBa. Statistics regarding the mental health scores are before standardization within the analytic sample. Intra-class correlation is between parental siblings.

Observational associations

We ran a linear regression to estimate observational associations between parental symptoms and children academic achievement. In the total MoBa sample, most parental mental health scores are negatively associated with their children's academic achievement (Figure 2 – black and Supplementary Tables 5 & 6). All mental health scores were negatively associated with scores in mathematics and reading tests, with standardized βs from -0.025 for ADHD (SE = 0.00) to -0.058 for eating disorders at age 8 (SE = 0.00), expect postnatal alcohol problematic use (β = 0.00, SE = 0.00 & β = 0.01, SE = 0.00). Associations of anxious and depressive, ADHD and eating disorders symptoms are significant but relatively weak: an increase of one standard deviation in the parental symptoms score is associated with

a decrease of maximum 0.06 SD of the child's standardized test score. For comparison, being a boy was associated with an increase of 0.17 SD in maths and 0.11 in English, and a decrease of 0.12 SD in reading. Surprisingly, higher English tests scores were associated with higher parental anxious and depressive symptoms in pregnancy (β = 0.009, *SE* = 0.003) and at child's age 8 (β = 0.009, *SE* = 0.004), and with postnatal alcohol problematic use (β = 0.02, *SE* = 0.004). Parental ADHD symptoms are not associated with English tests scores (β = -0.006, *SE* = 0.004), and eating disorders symptoms were negatively associated (β = -0.026, *SE* = 0.004 for symptoms during pregnancy, and β = -0.012, *SE* = 0.005 for child's age 8 symptoms), as expected and found for maths and reading scores.

We ran the same regressions in the children-of-siblings subsamples, to investigate the specificity of this subset of the MoBa participants, and enable a comparison with a within-sibling model in the exact same subset. As expected given the smaller sample size, the estimates are less precise in the children-of-siblings subsamples (Figure 2 – blue & Supplementary Tables 2 & 3). Overall, the associations are similar, but some point estimates are lower and many are no longer significantly different from zero. Reading tests scores are only significantly associated with parental eating disorders symptoms in this subsample, and English tests scores with anxious and depressive symptoms in pregnancy. Anxious and depressive and eating disorders symptoms are still significantly associated with lower maths scores in this subsample.

Within-sibling associations

After controlling for shared familial factors with a children-of-siblings design, no association between parental symptoms and children's test scores was statistically significant after multiple testing correction (Figure 2 – green and Supplementary Tables 2 & 3). Few within-sibling associations were suggestively significant (by suggestively significant we mean significant only without correcting for multiple testing, i.e. p-value < 0.05 without false-rate discovery correction). Children's maths tests scores were suggestively associated with parental anxious and depressive symptoms during pregnancy (β = -0.025, *SE* = 0.010) and at age 8 (β = -0.044, *SE* = 0.020), and with eating disorders symptoms at age 8 (β = -0.060, *SE* = 0.028). Children's reading and English tests scores were suggestively associated with eating disorder symptoms at age 8 (β = -0.058, *SE* = 0.028 and β = -0.058, *SE* = 0.029 respectively). For eating disorders symptoms, controlling for shared familial factors by comparing children-of-siblings led to a reduction in the point estimates of the associations for symptoms during pregnancy. However, the point estimates with symptoms at age 8 was not reduced (the estimates are imprecise nevertheless.

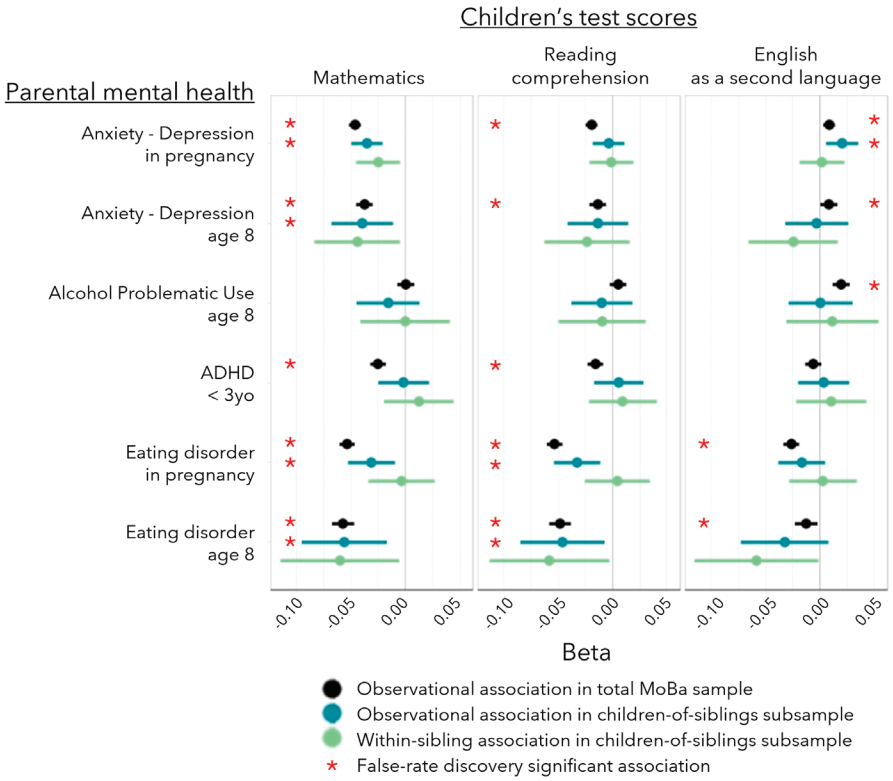


Figure 2 | Association of parental mental health scores with children's academic achievement. Coefficients as estimated with a linear regression in the total MoBa samples (black), in the children-of-siblings subsamples (blue), and with the children-of-sibling model (green). Red asterisks indicate estimates that reach statistical significance ($p < 0.05$) after false-rate discovery correction for multiple testing. Bars are 95% CIs. Child ages at which the parent reported their mental health are approximate ages (see Methods).

We also ran all children-of-siblings analyses in same-sex sibships, estimates are similar to those presented here: none of the associations reach suggestive significance (Supplementary Tables 7-9 and Supplementary Figure 2). Additionally there is no visual evidence for non-linear relationship when plotting individual data and quartiles trends.

DISCUSSION

Ten-year-old children (5th grade) tended to score lower on standardized tests of mathematics and reading if their parents had more symptoms of anxiety, depression, ADHD, or eating disorders. After controlling for shared familial factors by comparing up to 17,000 children of siblings, all significant associations disappeared, which suggests that parental mental health does not impact children's academic achievement. Some associations might remain

statistically significant in a larger children-of-siblings sample, but our study suggests an increase of one SD in parental mental health scores would decrease their children's test score by 0.06 SD at most.

Our study looks at the effect of parental self-reported symptoms, in a population sample of Norwegian parents. Exposures to more severe parental psychopathology might have larger effect on children's test scores. Indeed, besides exposure to their parents' symptoms, severe psychopathologies might result in additional adversity for the children, such as parental divorce and absence of parents due to hospitalization[311–313]. However, our conclusions based on self-reported symptoms are supported by a recent study in the full (Norwegian) population using diagnosed internalising disorders (Nordmo et al., in prep). Nordmo et al. also use a genetically-informed design, comparing test scores of 16-year-old siblings differentially exposed to their parents' internalizing disorders. They find that only exposure to parental psychopathology close to academic testing appears to have an effect, of small size.

Parental symptoms experienced close in time to the child's academic test may matter more than parental symptoms during pregnancy. Children born to mothers with an eating disorder during pregnancy had on average lower test scores across domains (β from -0.06 to -0.02), but these effects completely disappeared ($\beta \sim 0$) after controlling for shared familial factors, which suggests that shared factors cause the association. In contrast, we do not observe this reduction for maternal eating disorder symptoms when children were 8: After controlling for shared familial factors, parental symptoms stay suggestively associated with lower maths' achievement scores two years later (βs ~ 0.06). A similar pattern, though less pronounced, is visible for anxious and depressive symptoms. This timing effect awaits confirmation using other disorders, cohorts, and designs.

We find that the association of children's test scores with parental mental health depends on the type of symptoms. However, as we do not find evidence for a causal effect, these different associations are most likely due to different symptoms relating differently to shared familial factors, such as socioeconomic status or shared genetics. For example, parents who score higher than average on ADHD symptoms tend to have children who also score higher on ADHD symptoms. Previous research has shown that this is mostly due to the transmission of genetic liability, rather than the effect of being exposed to parental ADHD[157,314,315]. As childhood ADHD generally decreases academic performance[316], children of parents with ADHD are likely at risk for poor educational outcomes not because of parental exposure, but because of their own genetically-transmitted ADHD symptoms. Similarly, anxiety is thought to be transmitted to their children mostly by socialisation[317,318], children of anxious parents are more likely to have lower scores due to environmental factors shared with their parents,

than to shared genetics. Alternatively, we can not rule out that how we measure different aspects of parental mental health (some disorders could be better captured than others) might affect our comparison of parental disorders.

A surprising result is that the associations differ for maths and reading versus English as an additional language. While maths and reading skills are consistently negatively associated with parental mental health (apart from alcohol problematic use), English is surprisingly positively associated with parental anxiety, depression, and especially problematic alcohol use. We can speculate that this positive association is created as both alcohol use disorders and children' screen usage (facilitating home exposure to English) are more prevalent among lower-educated families[319,320]. Conversely, it could also be created as both alcohol consumption and active exposure to English are more common in higher-educated families[321–323]. Or again, this positive association could be an artifact as we use self-report problematic alcohol use, and people with alcohol use disorders might participate less in Moba[324]. In any case, the association disappeared in the within-family analysis, suggesting that the observed link is not causal.

There are limitations to our study and its interpretation. First, despite a big sample size, our sample is not fully representative of the Norwegian population. This is evident from the mean standardized test scores in MoBa, which are 0.11–0.27 SD above the population mean, and that in the children-of-siblings subsample, which are 0.11–0.34 SD above the population mean. Moreover, we see in **Figure 2** that the full MoBa and children-of-sibling subsample sometimes give different point estimates of the observational associations; see for example parental ADHD symptoms (but note the confidence intervals overlap). Ascertainment bias might be a concern, as parents with mental disorder or with lower SES are less likely to participate in surveys[325]. We recommend replication in an even bigger, more representative sample. Second, our children-of-sibling design is also subject to the same limitations as typical sibling comparison designs. Random measurement errors of the parental symptoms might accentuate the attenuation of the within-sibling effect[326]. How within-sibling designs are affected by ascertainment bias depends on the factors leading to selection. If those are stable and shared across siblings, they will be adjusted for by design[327], but if those are not shared within siblings, a within-sibling design might amplify the bias due to these factors[328]. Third, we stress that our study's conclusions are applicable to the Norwegian context. Norway is a country with a Scandinavian social-democratic welfare state system[329]. Our context is thus one where effects of parents' psychopathologies can be expected to be relatively weak in comparison with other contexts such as the U.S. or the U.K. Our finding of no impact of parents' psychopathologies on children's educational achievement in

Norway may not extend societies with less extensive welfare states. Finally, we recommend investigating whether exposure to parental psychopathology impacts offspring's educational outcomes later on, up to their highest obtained degree. Our observed parental effects at age 10 that seem negligible may accumulate, or parental effects may be larger in secondary and tertiary education[157,272].

Conclusion

While our study has limitations that prevent us from firmly establishing a causal relationship, our findings suggest that parents' mental health symptoms are unlikely to have a significant impact on their children's academic performance. We estimate that such effects, if they exist, are likely to be small, explaining little of the variation among children in their standardized test scores. While our results do not diminish the importance of addressing mental health concerns in parents, they do suggest that improving parental mental health may not in itself lead to improvements in children's educational outcomes.

CODE AND DATA AVAILABILITY

All code associated with the analyses is available on GitHub at https://github.com/PerlineDemange/ParentMH_childEA_Moba. MoBa data are available to individuals who obtain the necessary permissions from the data access committee (see https://www.fhi.no/en/studies/moba/for-forskere-artikler/research-and-data-access/).

CHAPTER 6

SUMMARY OF MY RESEARCH CHAPTERS

Education and mental health shape individuals' wellbeing and their opportunities in life. Understanding the relationship between these two domains and their intergenerational transmission is valuable for researchers, policy makers, and society as a whole. With this doctoral dissertation, I tried to further our understanding of the causes of individual differences in educational outcomes, their relationship with mental health, and their co-intergenerational transmission. In my research I worked with observational data in large samples, triangulating across several genetically-informed designs to study genetic and environmental transmissions. In this chapter, I summarize the main findings of my four empirical chapters.

CHAPTER 2: WHAT ARE NONCOGNITIVE SKILLS AND DO THEY MATTER?

Takeaway: Noncognitive aspects of educational attainment matter: they are estimated to account for 57% of the genetic variance in educational attainment and are genetically associated with other life outcomes to the same extent as cognitive skills. Noncognitive skills are a multifaceted construct: our noncognitive factor of educational attainment correlates genetically strongly, but not uniquely, with several personality and behavioural traits. The biology of the noncognitive factor appears very similar to those of the cognitive factor of educational attainment.

In **Chapter 2,** I leveraged genomic structural equation modelling to disentangle the cognitive and noncognitive components of education attainment. While noncognitive skills are a widely used concept, what constitutes the set of noncognitive skills is not widely agreed upon, and there is a lack of consistent measurements in large cohorts[30]. To overcome these challenges, our study was designed to mirror the original conceptualization of noncognitive skills as all traits positively contributing to educational success that are not cognitive skills[31]. I applied a new statistical approach which we named GWAS-by-subtraction: I identify genetic associations with a latent noncognitive trait (NonCog) by 'subtracting' genetic influence on cognitive performance (Cog) from the association of each genetic variant with educational attainment (EA). I then conducted phenotypic and biological annotation analyses to explore this genetic noncognitive construct, using LDSC genetic correlations, polygenic score (PGS) predictions, and enrichment analyses.

I successfully identified a genetic component of educational attainment that is independent of cognitive skills. In five cohorts, NonCog PGS predicted academic skills and its relationship to IQ scores was attenuated relative to Cog PGS. However, the correlations were not

attenuated to zero. We suggest this is due to the cognitive performance GWAS not capturing all forms of cognitive skills, and therefore its subtraction from EA leaves a residual cognitive signal. Despite a simplified and incomplete statistical separation of NonCog from Cog, we argue our NonCog GWAS is a useful tool to explore EA and noncognitive traits.

NonCog matters for education and life outcomes: the noncognitive component accounts for 57% of the genetic variance in educational attainment. The noncognitive component was associated with SES-related traits (income, neighbourhood deprivation) to a similar degree as the cognitive component. These results contribute new evidence that heritable individual differences in traits other than intelligence influence educational attainment and downstream life outcomes.

Phenotypic annotation analyses helped to gain a deeper understanding of the substance of heritable noncognitive aspects of education attainment and highlight their diversity. The genetics of NonCog are correlated with a combination of personality traits that resemble those that emerge during maturation into adulthood[330]: higher levels of openness to experience, conscientiousness, agreeableness, extraversion, and lower levels of neuroticism. In addition, NonCog genetic factors are correlated with "mature" decision-making preferences[331,332] such as lower risk-taking and lower present-oriented time preference. Consistently, NonCog genetic factors were associated with lower health-risk behaviours and later fertility. These associations highlight that the noncognitive component of educational attainment is unlikely to be a singular construct. No single "noncognitive" construct is solely responsible for the variance in educational attainment beyond cognitive skills. While NonCog seems associated mostly with socially desirable traits, it is also correlated with a higher risk for several mental disorders. These findings caution against an assumption that genetic variants associated with achieving higher levels of formal education are always associated with positive outcomes. This underscores the importance of understanding the complexity of the genetic variants that influence educational attainment and the potential implications for mental health.

Although this study focuses on the differences between noncognitive and cognitive skills in terms of their phenotypic annotation, the biological annotation reveals that these components may not be as different as we initially thought, despite being designed not to be genetically correlated. Our findings indicate that NonCog genetics are enriched in the same brain tissues and cell types as Cog. Although the genetic correlations between brain volumes and NonCog were different from those with Cog (Cog seemingly more associated with grey matter and NonCog with white matter), these correlations need to be replicated in future studies, as the neuroimaging GWASs we had access to were relatively low powered. On the other hand, the low differentiation at the cellular level suggests that both types of

skills likely involve similar biological processes.

In conclusion, by conducting a GWAS of a phenotype that was not directly measured, I offer a view of the genetic architecture of the noncognitive aspects of educational attainment. My results demonstrate that noncognitive skills are central to the heritability of educational attainment and establish connections between the genetics of these skills and of other social and behavioural traits.

CHAPTER 3: DO PARENTS' NONCOGNITIVE SKILLS AFFECT THEIR OFFSPRING'S EDUCATIONAL OUTCOMES?

> **Takeaway:** By combining data from three cohorts and three designs for estimating indirect genetic effects, I provide evidence for the environmental effects of parents' characteristics associated with cognitive and noncognitive skills genetics on offspring educational outcomes. I also performed extensive simulations which highlight subtle differences between the designs.

In **Chapter 3,** multiple genetic approaches were used to study the environmental effects of parents' skills on offspring education. I computed polygenic scores based on the GWASs of cognitive and noncognitive aspects of educational attainment (**Chapter 2**) in three different cohorts. I estimated their associations with educational achievement and attainment. Via three different family-based designs, i.e. comparing siblings from the same family, adoptees, and parent-offspring trios, I could disentangle the genetic effect of the offspring's genotype on the offspring's education (direct genetic effect), and an environmentally mediated effect of the parental genotype on the offspring's education (parental indirect genetic effect). These designs estimate the environmental effects of polygenic scores (PGS) by estimating the effects of parental genetic variants not transmitted to the offspring (and therefore acting via the offspring's environment), by contrasting PGS estimates for adopted and non-adopted children, and by contrasting PGS effects obtained with and without comparing siblings.

I found evidence that parental characteristics tagged by both NonCog and Cog polygenic scores are associated with offspring education. Indirect genetic mechanisms explained 36% of the effect of the NonCog PGS and 40% of the effect of the Cog PGS. This demonstrates the presence of genuine environmental influences (associated with parents' genes), unconfounded by offspring-led effects of inherited genes. Further studies should investigate the mediating parental characteristics. These heritable environmental influences might be proximal (e.g. parenting) or more distal (e.g. the neighbourhood parents live in, a set of cultural practices).

The environmental effects of parents on offspring education tagged by NonCog and Cog PGS were consistent across countries, generations, outcomes, and analytic designs, with two exceptions. Estimated parental indirect genetic effects were null for childhood achievement in a Dutch cohort (NTR) but not for comparable outcomes in a UK cohort (TEDS). Parental indirect genetic effects estimated with the adoption design were lower than for the sibling and non-transmitted PGS designs.

Simulations provided further understanding of differences between the three statistical designs and their sensitivity to different components and biases. The adoption-based estimates of indirect genetic effects do not account for prenatal effects and appear more robust to population stratification and assortative mating. The sibling design seems particularly affected by sibling indirect genetic effects. Sensitivity analyses in the cohorts suggest the potential presence of population stratification, especially in the case of NonCog genetics, and of some assortative mating, but I find no evidence of sibling indirect genetic effects. This suggests the adoption design provides a lower-bound estimate of indirect genetic effects.

In conclusion, combining three cohorts and three designs for estimating indirect genetic effects, I provide evidence for environmental effects of parents' characteristics associated with cognitive and noncognitive skills genetics on offspring educational outcomes.

CHAPTER 4: IS THE RELATIONSHIP BETWEEN EDUCATIONAL ATTAINMENT AND MENTAL HEALTH CAUSAL?

> **Takeaway:** In a within-sibling design in the Dutch national registry (CBS) and using two-sample mendelian randomization, I found potential causal effects of educational attainment on the risk of being diagnosed with major depressive disorder (MDD), attention deficit hyperactivity disorder (ADHD), post-traumatic stress disorder (PTSD), alcohol dependence, and generalized anxiety disorder (GAD), as well as a causal effect in the opposite direction for ADHD. Results were inconsistent for schizophrenia, obsessive-compulsive disorder (OCD), bipolar disorder, and anorexia nervosa, suggesting different relationships between EA and the diagnosis liability, and EA and the diagnosis itself.

In **Chapter 4**, I evaluate the causal relationship between educational attainment and mental health using two quasi-experimental research designs. Firstly, I analyse Dutch national registry data to evaluate the causal association of education attainment with the risk of being diagnosed with one of 17 psychiatric disorders. To do so, I compare siblings' education

and diagnoses, hence fully controlling for unmeasured familial confounders such as shared genetic risk and family SES. Secondly, I apply two-sample mendelian randomization to study the relationship between education and 9 psychiatric diagnoses, based on summary statistics from numerous large GWAS studies.

Triangulating across these two methods, our results suggest that higher education attainment reduces the risk of MDD, PTSD, alcohol dependence, ADHD, and GAD diagnoses. I also find evidence for reverse causation where ADHD, and suggestively MDD, PTSD and GAD, influence educational attainment. Some of our results highlight the need for triangulation as inconsistent results were observed for schizophrenia, OCD, bipolar disorder, and anorexia. Within-family analyses suggest higher education reduces the risk of being diagnosed, while mendelian randomization suggests more education does not affect this risk for schizophrenia, and actually increases this risk for OCD, bipolar disorder, and anorexia nervosa.

Based on these inconsistent results, we hypothesise that these last diagnoses follow a model where the (genetic) liability for being diagnosed is associated with higher education, while the disorders themselves interfere with schooling. Supporting this, in the Dutch register data I found a high prevalence of bipolar disorder among pre-university high-school dropouts, and healthy siblings of bipolar and anorexic patients have a similar or higher average educational attainment than siblings from unaffected sibships. Further research is needed to disentangle the effects of genetic liability and disorder symptoms. Potential sources of bias need to be explored, such as inequalities in healthcare access, preferential diagnosis, or selection bias in the GWAS on which we based our mendelian randomization analyses (henceforth not capturing the true genetic liability of these disorders).

In conclusion, using within-family analyses and two-sample mendelian randomization, these results support that higher education attainment reduces the risk of certain psychiatric disorders, while some disorders can also influence educational attainment. However, some findings were inconsistent, highlighting the importance of triangulation. Further research is needed to disentangle the effects of genetic liability and disorder symptoms.

CHAPTER 5: DO PARENTS' MENTAL HEALTH AFFECT THEIR CHILDREN'S EDUCATIONAL OUTCOMES?

Takeaway: I find weak associations between parental psychopathology symptoms and child educational achievement at age 10 in Norway. When controlling for shared factors among siblings in the parent generation in a within-sibling design, the effects were further attenuated and no longer significant.

In **Chapter 5**, I compare up to 9,000 families of the Norwegian Mother, Father, and Child Study (MoBa). I selected families in which two of the parents are siblings (i.e. their offspring are cousins). This design allows for estimating the association between parental mental health and children's academic achievement, while controlling for unmeasured factors shared among adult siblings (e.g. genetics and socioeconomic status) that might confound the relationship.

The results suggest that parents' mental health symptoms are unlikely to have a significant impact on their children's academic performance. I found that children tended to score lower on 5th-grade standardized tests of mathematics and reading if their parents had more symptoms of anxiety, depression, ADHD, or eating disorders. However, after controlling for shared familial factors by comparing children who were cousins, all significant associations disappeared.

In conclusion, our findings suggest that any potential effects, if they exist, are likely to be minimal. Although addressing mental health concerns in parents is important, our results suggest that improving parental mental health alone may not result in better educational outcomes for their children.

CONCLUSION

My research has contributed to our understanding of education, mental health and their relationships. First, it showed that noncognitive aspects of educational attainment account for 57% of the genetic variance in educational attainment and reflect several personality and behavioural traits. Second, my research demonstrated that parental characteristics genetically associated with both cognitive and noncognitive aspects of educational attainment affect their offspring's education, above genetic transmission. Third, the research revealed causal effects of educational attainment on several psychiatric diagnoses, and suggest further mechanisms explaining surprising education-diagnoses associations. Finally, parents' mental health symptoms are unlikely to significantly impact their children's academic achievement in Norway. In the next chapter, I provide a reflection on my work.

07

CHAPTER 7

GENERAL DISCUSSION AND FUTURE DIRECTIONS

In this discussion, I consider the broader implications of my research, its value for the field and, briefly, how others have built upon it. I cover the benefits and challenges that arise from re-using observational data. I reflect on the value of triangulation for my work. I highlight the usefulness of my operationalization of noncognitive skills and suggest further examination of noncognitive skills as a missing link between education and mental health. I consider the nature of indirect genetic effects, which have become a cornerstone of intergenerational research based on polygenic scores. Finally, I discuss the challenges of unmodelled sibling effects and limited external validity for the future of the field.

OPTIMAL (RE)USE OF OBSERVATIONAL DATA

In this doctoral dissertation, I leverage a diverse range of existing observational data sources. I use population registry data from Statistics Netherlands (**Chapter 4**) and Statistics Norway (**Chapter 5**). I analyse data from diverse studies, which include among others longitudinal, survey, cohort, and biomedical data, and have diverse collection procedures. I also work with summary statistics of genome-wide association studies (**Chapters 2, 3, & 4**) and other publicly available gene expression and gene-set data (**Chapter 2**).

The existence and availability of these data resources enable extensive and valuable secondary analyses to be conducted without the need to collect new data. Secondary data analysis is resource-efficient but comes with challenges, whose possible solutions I illustrate in this doctoral dissertation.

One challenge is the need to accommodate existing measures rather than rely on the "perfect" measure. This requires accommodating proxy measures and alternative operationalizations. Our conceptualization of noncognitive skills in **Chapter 2** is one such example: instead of collecting measures of noncognitive skills in large samples, I leveraged available summary statistics of two GWASs. This was enabled by the new method Genomic-SEM, which allowed to combine data from multiple existing GWASs jointly to model noncognitive skills as a latent variable.

Another challenge is the impossibility to design retrospective randomized experiments to infer causal relationships. However, insofar as natural experiments occur in the data these can offer a suitable and powerful alternative. Here I leverage numerous natural experiments and quasi-experimental methods such as family-based designs (including within-sibship design and adoption design) and mendelian randomization. The difference we observe between observational associations (obtained from simple linear regressions controlling selected covariates) and within-sibship associations (controlled for all factors shared between siblings) in **Chapter 4** and **Chapter 5** is a clear illustration of the value of

such quasi-experimental designs. For example, while educational attainment has an equal association with ADHD and MDD diagnoses (OR = 0.9), controlling for all environmental and genetic influences shared between siblings suggests an additional year of education lowers the odds of an ADHD diagnosis by 9% but only by 3% for MDD. Mendelian randomisation confirms a similarly attenuated effect of education on MDD.

I relied on a range of invaluable resources throughout my research. One such resource was the Dutch population registry (CBS), which provided access to a substantial portion of the population and allowed for the identification of siblings with data on mental health and educational outcomes. I further relied on a large (N ~500.000) biomedical research cohort and several large (N ~2.000-20.000) longitudinal, twin, family, and developmental cohorts, with genotyped participants. These large datasets mostly provided ample statistical power, such that a more prominent issue was the potential for biases. I focused on countering bias by triangulating across various types of natural experiments that I can leverage in the different resources.

Finally, beyond the immediate use of cohorts or population registries in isolation, the linkage of national registry data with survey data and genotyped cohorts allows for very promising studies. In **Chapter 5**, standardized scores of MoBa children were obtained from the National Educational database, allowing to standardize the scores on the entire Norwegian population and access to educational outcomes directly without missingness. The Norwegian registry data further provided me with pedigree information of MoBa participants, allowing me to identify families in which two of the parents are siblings in MoBa while analysing the psychiatric symptoms they reported in MoBa surveys. Future studies could expand on these integrated data and further identify extended pedigree (their cousins, their sibling-in-law, a new partner, etc.) and combine it with available genetic data.

THE VALUE OF TRIANGULATION

Triangulating refers to the use of multiple methods -and samples- with their own set of assumptions and limitations to study a single phenomenon (here the intergenerational transmission of education and mental health). It is a means to avoid potential biases arising from the use of a single methodology/sample. The convergence of findings strengthens our confidence, while inconsistent or contradictory findings raise scepticism but can also suggest new interpretations of the phenomenon[13,15]. In order to properly leverage the advantages of genetically-informed designs while mitigating their limitations, I aspire to triangulate across different methods and different samples.

In my doctoral dissertation, triangulation allowed me to gain insights that would have been

otherwise unattainable. An instance of this is highlighted in **Chapter 2**, where both genetic correlations using out-of-sample GWAS summary and meta-analysed polygenic scores (PGS) predictions across six cohorts from three different countries revealed that the noncognitive factor is as (or more) strongly associated with educational attainment as the cognitive factor, but significantly less associated with cognitive test performances. This convergence strengthens our confidence in this finding. In **Chapter 5**, I observe no association of self-reported parental symptoms with their children's academic achievement, in a non-clinical sample. However, Nordmo et al.[333] (in a paper I co-author) find consistent null-to-small effect using diagnoses of internalizing disorders in the Norwegian national registry data. Both studies' findings are strengthened by these similar findings using alternative samples and measurements of mental health, with different limitations[334].

Inconsistent findings also highlight that employing only one design might have led to incorrect, or at best incomplete, conclusions. For example, in **Chapter 3**, the use of different genetic designs revealed lower indirect genetic effect estimates for educational attainment when using an adoption design compared to a within-sibship or non-transmitted PGS design. Simulations and understanding of the biases inherent in each design led us to hypothesise there might be prenatal indirect genetic effects, population stratification, and/or assortative mating. Having identified plausible causes of the differences, these could then be tested, and we conclude there likely is population stratification and potentially prenatal effects. Similarly, in **Chapter 4**, findings for bipolar disorder, anorexia, and OCD differed in sign between the mendelian randomization and within-sibship design. As both methods differ in the underlying construct they capture (genetic liability vs diagnosis), this suggested potential differences in causal effects of the genetic liability and the diagnosis of the same disorder. These observations helped to formulate new hypotheses that can be further tested.

A crucial point to consider is that triangulation across different designs and samples cannot remedy a flawed study. It is necessary to properly interpret the results taking into account the assumptions, potential limitations, and biases of the study. By adopting a triangulation approach, researchers are compelled to explicitly state the underlying assumptions of their model. Combined with pre-registration, this encourages a more thorough evaluation of the study's added value in addressing the research question. Additional to triangulation, it is essential to conduct research using state-of-the-art approaches to maximizing reproducibility and reliability, while also maintaining epistemic humility when drawing conclusions[335].

A SIMPLE BUT USEFUL MODEL FOR NONCOGNITIVE ASPECTS OF EDUCATIONAL ATTAINMENT

Leveraging genome-wide associations studies and Genomic-SEM in **Chapter 2**, I could measure the association of genetic variants with noncognitive skills, while operationalizing noncognitive skills as a residual of the subtraction of cognitive skills from educational attainment (EA). This operationalization was necessary to counteract the lack of consistent measures of noncognitive skills and make this study possible.

Separating noncognitive from cognitive aspects of EA is an approach that, while reductive, proved a highly effective first-order approximation of what is fundamentally a nuanced and complex developmental relation. The suggested dichotomy between cognitive and noncognitive skills is a well-known subject of criticism[336], as cognitive and noncognitive skills likely develop together, or in interaction with each other[123,124]. Similarly, we expect cognitive performance to not only affect educational attainment but also be affected by it[24].

In anticipation of these potential limitations, I performed various sensitivity analyses in **Chapter 2** and re-ran core analyses while introducing a positive correlation between cognitive (Cog) and noncognitive latent factors (NonCog), or while allowing for educational attainment to causally affect cognitive skills. Neither of these alterations significantly impacted the noncognitive factor suggesting the operationalization is robust to these alternate assumptions.

Ultimately a reductive conceptual model was necessary to render a GWAS of noncognitive skills possible. A form of reduction also remains necessary to facilitate the interpretation of a latent NonCog GWAS. Although our model of NonCog may be imperfect, its adoption by other researchers as a starting point for investigating noncognitive aspects of educational attainment underscores the value of our NonCog GWAS as a research tool.

Can we further separate noncognitive skills from cognitive skills?

Our GWAS-by-subtraction model aimed to isolate genetic variance in education that was independent of cognitive performance. Our GWAS of noncognitive aspects of educational attainment was designed as genetically independent of the GWAS of cognitive performance, with a genetic correlation of zero. However, in independent samples, NonCog had a steeply and significantly attenuated, albeit not zero, association with cognitive measures (IQ tests). The pattern of associations was replicated by independent researchers in the Brisbane Longitudinal Twin study[337], and in the context of new GWASs of language-related skills (word reading, nonword reading and nonword repetition)[338]. We suggest this imperfect statistical separation of NonCog from cognitive performance is due to the limitations of the cognitive

performance GWAS we had to rely on, which did not encompass a broad enough set of measures of cognitive skills. Consequently, its subtraction leaves a residual cognitive signal.

New research has extended our GWAS-by-subtraction to try to estimate a more specific GWAS of noncognitive aspects of educational attainment. Researchers extended our GWAS-by-subtraction replacing the GWAS of cognitive performance with a recent GWAS of intelligence based on more diverse cognitive measures[339,340]. Malanchini, Allegrini and colleagues (in a paper I co-author) included a range of cognitive traits (processing speed, executive functions, reaction time) and several socioeconomic indicators (neighbourhood deprivation, household income) in the model in order to capture noncognitive aspects of SES (not only EA) and to increase the differentiation between the NonCog and Cog factors[341]. These updates to the GWAS-by-subtraction result in a latent noncognitive factor with a genetic correlation larger than 0.95 with ours. The patterns of correlations with other traits were also largely consistent. This suggests the obtained NonCog genetics are virtually identical to ours, and these extensions do not meaningfully improve the statistical separation of NonCog from cognitive measures. This somewhat surprising result raises the question of whether a stricter statistical separation of NonCog from cognitive measures is even possible.

What are heritable noncognitive aspects of educational attainment?

Phenotypic annotation analyses in **Chapter 2** show consistent genetic associations of NonCog with mature personality traits and mature decision-making preferences, with lower health-risk behaviours, and with later fertility. These associations highlight that the noncognitive component of educational attainment is a multifaceted construct, of generally socially desirable characteristics (aside from specific positive associations with psychiatric disorders).

Other uses of our available Cog and NonCog summary statistics help us complete the picture. Notably, Abdellaoui[342] and colleagues ran genetic correlations of Cog and NonCog with numerous traits, cardiovascular and physical health traits among others. Additional to the differences we report, the most salient differences between Cog and NonCog genetic correlations reported by Abdellaoui et al. confirm that NonCog is particularly positively associated with behaviours improving health and related health outcomes: NonCog is less associated with exceeding motorway speed than Cog, positively associated with moderate and intense physical activity and sleep duration (while Cog is negatively associated), and negatively associated with diabetes (which has no association with Cog).

My conceptualization of NonCog as “everything which affects educational attainment but cognitive performance” has the benefit to give access to the overall heritable noncognitive

aspects of education attainment. This allows for confirmation of an overall effect before investigating specifics. In **Chapter 3**, I therefore can quantify the overall environmental effects of parents on offspring education tagged by NonCog polygenic scores. As these indirect genetic effects are non-null, follow-up studies to investigate mediators might be of interest. Parental characteristics to prioritize in this follow-up investigation are traits highly associated with NonCog genetics. For example, we suggested future research could investigate parental depression. **Chapter 5**'s insignificant effects of parental mental health symptoms on their children's academic achievement yet suggest that indirect genetic effects are unlikely to be mediated by parental mental health in Norway.

When contrasting our findings, with the literature on noncognitive skills[30,343–345], it is noteworthy that noncognitive skills were sometimes measured (and/or conceptualized) as the results of behaviour screening tools developed by psychologists and psychiatrists to identify children's externalizing and internalizing problems[346,347]. Among these are the Strengths and Difficulties Questionnaires (SDQ)[348] and the Child Behaviour Checklist (CBCL)[349]. However, a NonCog PGS based on our GWAS was reported to not be associated with measures of internalizing behaviour in ALSPAC[350,351], measured with SDQ. In adulthood, NonCog PGS was not associated with neuroticism in NTR, Texas Twins and WLS cohorts (**Chapter 2**) and the genetic correlation of NonCog with neuroticism, anxiety and worry is potentially null[339,340]. Overall, these results imply that these behaviour screening tools might not be the most appropriate measure of heritable noncognitive aspects of educational attainment. This is a good example where our GWAS of a latent construct can guide the selection of measures of observed noncognitive skills.

GWAS-BY-SUBTRACTION OF NONCOGNITIVE SKILLS: A FURTHER TOOL TO UNDERSTAND MENTAL HEALTH?

In **Chapter 2**, I estimated the genetic correlations of cognitive and noncognitive aspects of educational attainment with several psychiatric disorders. This analysis revealed that NonCog genetic factors are associated with a higher risk for multiple psychiatric disorders: anorexia nervosa, OCD, bipolar disorder and schizophrenia. These findings have three key implications. First, the positive association of educational attainment with some psychiatric disorders seems paradoxical and is mostly driven by its noncognitive component. Second, these results support the hypothesis that psychiatric disorders may reflect extreme expressions of psychological traits that are also present within the normal range of adaptive functioning. Third, investigating the comparison with Cog and NonCog might be crucial in understanding the link with education, and in differentiating psychiatric disorders.

Noncognitive aspects of educational attainment help to understand the link between education and mental health

Investigating the role of cognitive and noncognitive skills in mental health is therefore a promising future research direction. Fortunately, other researchers already started to tackle it by leveraging the tools I developed in **Chapter 2**. Several studies have looked at additional correlations between Cog and NonCog and psychiatric traits[337,339,340,352]. In **Figure 1**, I combined all available published genetic correlations to date. Most psychopathologies are more associated with NonCog than Cog (positively or negatively), with the notable exception of anxiety (depending on the GWAS used), autism spectrum disorder, and Alzheimer's disease for which NonCog shows no association.

Psychiatric disorders for which I found a risk-increasing effect of educational attainment with **Chapter 4**'s mendelian randomization (OCD, bipolar disorder, anorexia, autism) are all disorders to which NonCog is positively genetically correlated (**Figure 1**). For example, I found no effect of EA on schizophrenia from mendelian randomization approaches in **Chapter 4** and an opposite direction genetic association of Cog and NonCog with schizophrenia in **Chapter 2**. It would be valuable to extend our MR analyses to investigate if this null effect of EA on schizophrenia does translate to opposite directional effects of Cog and NonCog. Such analysis was done by Thorp et al.[339] in the case of Alzheimer's disease (AD). Using Cog and NonCog-associated genetic variants in a multivariate MR design, they show that the protective effect of education attainment on AD is exclusively due to the cognitive aspects of educational attainment. This finding has important implications for interventions aimed at preventing the development of AD through education, suggesting that interventions will not have a beneficial impact if they act through noncognitive rather than cognitive skills.

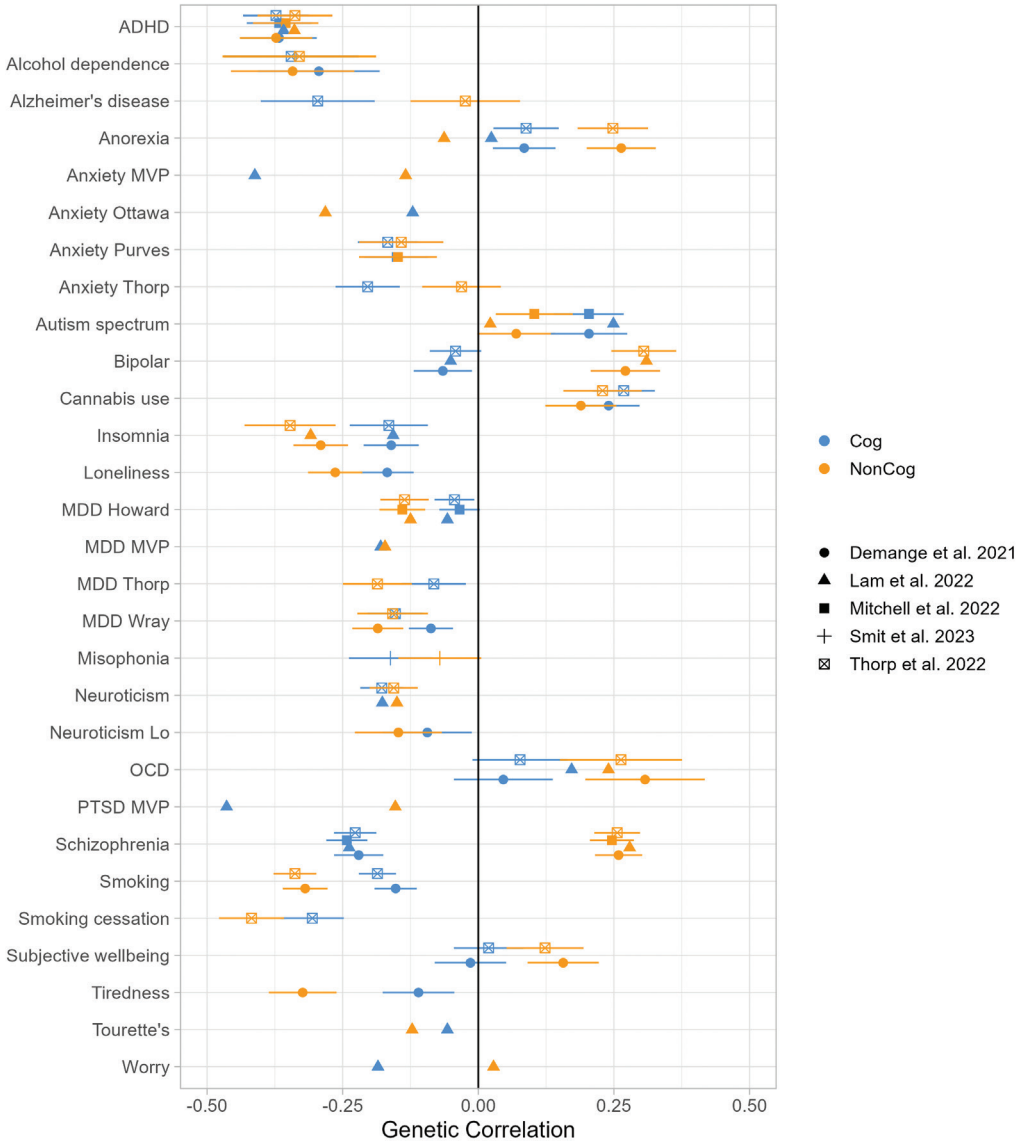


Figure 3 | Genetic correlations of Cog (blue) and NonCog (orange) with psychiatric disorders and related traits. These correlations were performed within 5 published manuscripts[152,337,339,340,352] (shape), estimates were obtained from available supplementary data (no standard errors are available for Lam et al.). Genetic correlations were estimated with LDSC, for specific methodology please refer to the original manuscripts. Psychiatric disorders are ranked by alphabetical order and followed by the name of the first author of the GWAS used in case of heterogenous results (MVP = Million Veterans program).

Noncognitive aspects of educational attainment and understanding the biology of psychiatric disorders

In **Chapter 2**, the biological annotation of Cog and NonCog surprisingly reveals that the same cell types mediate genetic influences on NonCog and Cog, which suggested that biological differentiation of Cog and NonCog would require finer-grained molecular data. Lam et al.[340] make an interesting step in this direction by using functional annotation of meta-loci (LD-independent regions showing similar local genetic correlation profiles across psychopathological traits). They found that genes associated both with Cog and psychiatric disorders featured genes predominantly involved with neurodevelopmental processes and were expressed prenatally, while NonCog meta-loci genes were expressed predominantly in early adulthood and adulthood. This is especially interesting in light of our own findings that noncognitive skills have a genetic association with personality traits reflecting adaptive adult behaviours which typically develop in early adulthood.

Some of these meta-loci uniquely correlate with specific psychiatric disorders and the annotation reveals interesting common pathways. For example, one of the meta-loci from Cog uniquely negatively associated with anorexia appears characterized by genes implicated in metabolism, suggesting metabolism as a common link between low cognition and anorexia. These fascinating results show the value of cognitive and noncognitive aspects in better understanding the biology of psychiatric disorders, as well as the possible underlying biological pathways linking them to educational attainment.

Using GWAS-by-subtraction to better understand psychopathologies?

The GWAS-by-subtraction approach we implemented in Genomic-SEM is applied by others as a tool to separate the genetic variance of two GWASs. It has been particularly used in genetic psychiatry to investigate the specificity of certain psychopathologies. For example, Ahangari and colleagues[353] separated the genetic factors unique to bipolar disorder and depression from the genetic factors shared with schizophrenia via GWAS-by-subtraction. Other applications include the investigation of the independent effects of neuroticism and depression on cardiovascular diseases[354], of wellbeing independent of depression disorder[355], and the conceptualization of cognitive resilience[356].

INDIRECT GENETIC EFFECTS: WHAT ARE WE REALLY LOOKING AT?

Indirect genetic effects are a powerful way to isolate environmental effects from genetic transmission. The work presented in **Chapter 3** was the first evaluation of indirect genetic effects of noncognitive skills. We were the first to compare three new methods using

polygenic scores to study these (in)direct genetic effects, which were gaining popularity quickly[170,285].

Importantly how we should interpret indirect genetic effects is still unclear. I tried to avoid referring to indirect genetic effects as proof of parental nurture in **Chapter 3** and I acknowledge that "the parental indirect genetic effects we have identified may capture proximal forms of 'nurture' […] and/or more distal environmental effects". However, I do conclude on "environmental effects of parents' noncognitive and cognitive skills on offspring educational outcomes", implying that parental skills are responsible for these indirect genetic effects. Nivard et al.[357] and Abdellaoui et al.[342] pointed out that indirect genetic effects have been mostly described as being genetic nurture "nature of nurture"[156], which implies a notion of nurturing process taking place within the nuclear family, such as parenting. On the other hand, others have rather interpreted indirect genetic effects as "dynastic effects", invoking a notion of intergenerational transmission of wealth, more distal, and eventually implicating process in the extended family[166,168].

Leveraging the large genotype and family data in the Norwegian MoBa cohort, Nivard and collaborators looked at siblings in the parental generation to try to disentangle genetic nurture from dynastic effects[357]. They regressed the child PGS with the average PGS of the sibling in the parental generation (between-sibling indirect effect) and the deviation of the parental PGS from the parental-siblings average (within-sibling indirect effect). This within-sibling indirect effect represents the indirect effect that operates through the nuclear family environment "genetic nurture", while the between-sibling indirect effect captures the broader "dynastic effects", capturing aspects of the environment shared in the extended family as well as previous generations' indirect genetic effects. They find that the between-family indirect genetic effect is significant while the within-family is not. This strongly suggests that the indirect genetic effects we observe should not be interpreted as genetic nurture, but as broader dynastic effects. Heritable parental skills might therefore not matter above the broader environment inherited from generations. This suggests characteristics tagged by NonCog and Cog PGS which environmentally affect their offspring's education are not proximal parental individual characteristics. This further partition of indirect genetic effects into "nurture" and "dynastic" aspects is interesting and invites caution in the interpretation of indirect genetic effects.

How the measure of indirect genetic effects might be differently captured by different designs is still ongoing research. In **Chapter 3**, our simulations show how population stratification, assortative mating, sibling indirect genetics and prenatal effects are differently accounted for by different PGS designs. These simulations could be extended to investigate how the

two “nurture” and “dynastic” aspects of indirect genetic effects are captured by the different family-based PGS designs, as well as how biases and components affect them. There are other methods to estimate indirect genetics effects other than PGS designs, how the estimates compare with each other still needs to be further clarified[163,182,183].

A potentially underestimated bias of PGS designs to estimate indirect genetic effects is that the construction of these polygenic scores is based on weights calculated from existing GWASs, which themselves could be biased by gene-environment correlation, population stratification, and assortative mating. This issue was highlighted for the sibling design[358], but is not yet investigated for other designs. One potential solution for this unknown bias is to rely on weights from a GWAS performed within-families. There are currently few available within-family GWASs[194] that are sufficiently statistically powered, but this is likely to change in the next years. Interestingly, Nivard et al. estimate the indirect genetic effects when using a PGS based on a within-sibling GWAS of education attainment. They observed that while the direct genetic effect does decrease, the indirect genetic effects are basically unchanged. This result should be replicated and explained.

FAMILY-BASED DESIGNS: BIAS DUE TO SIBLING EFFECTS MIGHT BE OVERLOOKED

Siblings can affect each other in ways that are complex and might be difficult to observe directly. Sibling effects can differ in their direction (positive or negative, cooperative or competitive) or on their symmetry (siblings affecting each other concurrently or only one sibling influencing another)[359,360], which can depend on additional sibling characteristics such as birth order or gender[361,362]. Moreover, the sibling effects do not have to originate from the siblings themselves, but can also result from parental behaviour, for example compensating or amplifying differences between siblings[363,364]. In part due to their complexity, sibling interactions are often ignored despite their potential confounding effects in all types of family-based designs.

Sibling effects might result in a violation of the Stable Unit Treatment Value Assumption (SUTVA), which assumes the exposure of one sibling to a risk factor does not influence the unexposed sibling[365–367]. In the context of **Chapter 4** for example, if the mental health of one sibling affects their education, which in turn affects the other sibling, this assumption does not hold. Sibling effects (including sibling indirect genetic effects) might therefore lead to important bias in family-based (PGS) designs, when considering causal effects. In **Chapter 3**, our simulations showed that the presence of sibling effects will bias estimates of parental indirect genetic effects with all three designs, and more strongly so in the case of within-

sibling PGS design. I therefore investigated the presence of sibling indirect genetics effects with three different methods in NTR and UK Biobank. We conclude that there are no sibling indirect effects of Cog and NonCog PGS on education in our samples. However, using different designs and a PGS of educational attainment in UK Biobank, Howe et al. found small sibling effects[184], contradicting our conclusion.

Veller and Coop emphasize that sibling interactions are a concern for the reliability of another key research design: within-sibling GWASs[368]. Within-sibling GWAS have been widely described as the most promising avenue to improve GWAS signals[369,370]. To mitigate potential bias arising from sibling effects, genome-wide associations should also be triangulated across several designs which would be affected differently, relying on adoptees or parent-offspring trios for example as we highlight in **Chapter 3**.

INTERNAL VALIDITY IS NOT ENOUGH: A WORD OF CAUTION ON EXTERNAL VALIDITY

Here I want to discuss the challenges of generalization of genetically-informed studies (external validity). These challenges are not unique to my research but deserve to be mentioned as resolving them is one of the goals for the next decade of (genetically-informed) epidemiological research.

When working with genomic data (**Chapter 2**, **Chapter 3**, and MR in **Chapter 4**), one key limitation of my work is the restriction to individuals with European ancestry. This choice was made for two reasons: most of the available genomic data are from individuals of European ancestry (representing 88% of all discoveries GWASs[371]), and genomic studies are not portable across populations due to population stratification[369]. As a result, it is difficult to conclude anything on the genetic architecture of noncognitive aspects of educational attainment and the causal effect of EA and MH as estimated with MR for individuals of non-European ancestry. Additionally, this restriction of my GWAS of NonCog (due to the European ancestry of the EA and cognitive performance GWASs it is based on) is allowing polygenic scores of NonCog to be estimated accurately with current tools only for European ancestries individuals, which perpetuate the exclusion of non-European ancestry individuals from these further studies. There is an urgent need for more diversity in genetic studies, solutions have been suggested elsewhere[372–374].

Asides from the lack of diversity in ancestry, all cohorts included in my four chapters are from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies[375] and self-selection of participants into these cohorts may further reduce the representativity of the cohorts. Selection bias might particularly be a concern for the traits studied, as individuals

with lower education and worse mental health are less likely to participate in studies[325,376]. Selection bias might moreover lead to collider bias and induce spurious associations[265]. Correction for the non-representativity of these cohorts can be partially done, for example weights for correcting the important volunteer bias of UK Biobank have recently been made available recently[263,264,377]. Similar efforts should be undertaken for all large cohorts on which most of the current genetically-informed studies are taking place. Crucially, methods based on family or using genetic data might induce a stronger selection, as multiple family members are required to volunteer, such is the case in **Chapter 5**. Using national registry data as I have done in **Chapter 4** solves many issues of data representativity, but it is still important to be aware of decisions around data collection and further study sample selection[334,378]. In my case for example, only selecting siblings might exclude some families, e.g. childless individuals and individuals who immigrated to the Netherlands whose siblings did not. Additionally, interpretation of findings in national registries still needs to be done in accordance with the measurement used: if a subset of the population is unlikely to be diagnosed (e.g. preferential diagnosis, inequalities in healthcare access, in health conceptualization[379]), my results will not pertain to this part of the Dutch population.

In sum, while the use of large cohorts in well-designed genetically-informed studies can lead to significant advancements, addressing the generalizability of the studies is a challenge and will be crucial to uncover actual mechanisms behind the aetiology and transmission of education and mental health. Together these examples show the value of triangulating not only across statistical designs but also across samples and measures.

AN EXCITING RESEARCH AVENUE

Family-based designs, genomic data, and their integration give us resources and tools to disentangle genetic and environmental transmissions and enhance causal inference. New methods are implemented to parse out genetic effects, environmental effects, gene-environment correlations and identify interactions. This fast-paced methodological development is opening opportunities to extend and refine designs used to investigate intergenerational transmission.

These methodological advancements go hand in hand with the increased availability of large longitudinal (familial) datasets with diverse sociological, economic, psychological, and biological measures, that are essential for such research designs. The opening of state biobanks, population registries, diverse cohorts, etc., offers exciting opportunities. Furthermore, the secure linkage of different sources of data and the success of large-scale collaborations are facilitated by (inter)national initiatives such as ODISSEI[380] or Tryggve[381].

The interdisciplinary research resulting from these developments offers new avenues for scientific progress. I have been privileged to be part of these thriving developments with this doctoral work. To paraphrase McAdams et al.[10]: It is an exciting time for research and I look forward to new developments yet to come.

REFERENCES

1. Hertz, T. *et al.* The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *BE J. Econ. Anal. Policy* **7**, (2008).
2. Shavit, Y. & Blossfeld, H.-P. Persistent Inequality: *Changing Educational Attainment in Thirteen Countries. Social Inequality Series.* (Westview Press, 5500 Central Avenue, Boulder, CO 80301-2847 ($45)., 1993).
3. Engzell, P. & Tropf, F. C. Heritability of education rises with intergenerational mobility. *Proc. Natl. Acad. Sci.* **116**, 25386–25388 (2019).
4. Lahtinen, H., Martikainen, P. & Tarkiainen, L. Changes in sibling similarity in education among Finnish cohorts born in 1950–89: the contribution of paternal and maternal education. *Longitud. Life Course Stud.* **13**, 496–526 (2022).
5. Galton, F. *Hereditary Genius: an Inquiry into its Laws and Consequences.* (MACMILLAN AND CO., 1892).
6. McLaughlin, K. A. *et al.* Parent psychopathology and offspring mental disorders: Results from the WHO World Mental Health Surveys. *Br. J. Psychiatry* **200**, 290–299 (2012).
7. Halpern-Manners, A., Schnabel, L., Hernandez, E. M., Silberg, J. L. & Eaves, L. J. The relationship between education and mental health: New evidence from a discordant twin study. *Soc. Forces* **95**, 107–131 (2016).
8. Morris, T. T. *et al.* Implications of the genomic revolution for education research and policy. *Br. Educ. Res. J.* **n/a**, (2022).
9. Pingault, J. B. *et al.* Using genetic data to strengthen causal inference in observational research. *Nat. Rev. Genet.* **19**, 566–580 (2018).
10. McAdams, T. A., Cheesman, R. & Ahmadzadeh, Y. I. Annual Research Review: Towards a deeper understanding of nature and nurture: combining family-based quasi-experimental methods with genomic data. *J. Child Psychol. Psychiatry* **64**, 693–707 (2023).
11. D'Onofrio, B. M., Lahey, B. B., Turkheimer, E. & Lichtenstein, P. Critical Need for Family-Based, Quasi-Experimental Designs in Integrating Genetic and Social Science Research. *Am. J. Public Health* **103**, S46–S55 (2013).
12. Harden, K. P. & Koellinger, P. D. Using genetics for social science. *Nat. Hum. Behav.* (2020) doi:10.1038/s41562-020-0862-5.
13. Lawlor, D. A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* **45**, 1866–1886 (2016).
14. Thurmond, V. A. The Point of Triangulation. *J. Nurs. Scholarsh.* **33**, 253–258 (2001).
15. Mathison, S. Why Triangulate? *Educ. Res.* **17**, 13–17 (1988).
16. Denzin, N. K. Triangulation. in T*he Blackwell Encyclopedia of Sociology* (John Wiley & Sons, Ltd, 2015). doi:10.1002/9781405165518.wbeost050.pub2.
17. Deary, I. J. & Johnson, W. Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *Int. J. Epidemiol.* **39**, 1362–1369 (2010).
18. Deary, I. J., Strand, S., Smith, P. & Fernandes, C. Intelligence and educational achievement.

Intelligence 35, 13–21 (2007).
19. Roth, B. *et al.* Intelligence and school grades: A meta-analysis. *Intelligence* **53**, 118–137 (2015).
20. Kevenaar, S., Dolan, C. V., Boomsma, D. & Bergen, E. van. Self-control and grit are associated with school performance mainly because of shared genetic effects. Preprint at https://doi.org/10.31234/osf.io/xt5qe (2023).
21. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
22. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat. Genet.* **49**, 1107–1112 (2017).
23. Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M. & Boomsma, D. I. Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Res. Off. J. Int. Soc. Twin Stud.* **5**, 544–553 (2002).
24. Ritchie, S. J. & Tucker-Drob, E. M. How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychol. Sci.* **29**, 1358–1369 (2018).
25. Smithers, L. G., Sawyer, A. C. P., Chittleborough, C. R. & Davies, N. Do early life " non-cognitive skills " matter ? A systematic review and meta-analysis of early life effects on academic achievement , psychosocial , language and cognitive , and health outcomes. *Nat. Hum. Behav.* **2**, 1–35 (2018).
26. Richardson, M., Abraham, C. & Bond, R. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychol. Bull.* **138**, 353–387 (2012).
27. Rimfeld, K., Kovas, Y., Dale, P. S. & Plomin, R. True grit and genetics: Predicting academic achievement from personality. *J. Pers. Soc. Psychol.* **111**, 780–789 (2016).
28. Heckman, J. J. & Kautz, T. Hard evidence on soft skills. *Labour Econ.* **19**, 451–464 (2012).
29. Domitrovich, C. E., Durlak, J. A., Staley, K. C. & Weissberg, R. P. Social-Emotional Competence: An Essential Factor for Promoting Positive Adjustment and Reducing Risk in School Children. *Child Dev.* **88**, 408–416 (2017).
30. Morris, T. T., Davey Smith, G., van den Berg, G. & Davies, N. M. Consistency of noncognitive skills and their relation to educational outcomes in a UK cohort. *Transl. Psychiatry* **11**, 1–10 (2021).
31. Bowles, S. & Gintis, H. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life.* (Routledge & Kegan Paul, 1976).
32. Heckman, J. J. & Rubinstein, Y. The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *Am. Econ. Rev.* **91**, 145–149 (2001).
33. Grusec, J. E. & Davidov, M. Parent Socialization and Children's Values. in *Handbook of Parenting* (Routledge, 2019).
34. England-Mason, G. & Gonzalez, A. Intervening to shape children's emotion regulation: A review of emotion socialization parenting programs for young children. *Emot. Wash. DC* **20**, 98–104

(2020).
35. Hart, S. A., Little, C. & van Bergen, E. Nurture might be nature: cautionary tales and proposed solutions. *Npj Sci. Learn.* **6**, 1–12 (2021).
36. Crossley, N. A. *et al.* The enduring gap in educational attainment in schizophrenia according to the past 50 years of published research: a systematic review and meta-analysis. *Lancet Psychiatry* **9**, 565–573 (2022).
37. Glahn, D. C., Bearden, C. E., Bowden, C. L. & Soares, J. C. Reduced educational attainment in bipolar disorder. J. *Affect. Disord.* **92**, 309–312 (2006).
38. Lorant, V. *et al.* Socioeconomic inequalities in depression: A meta-analysis. *Am. J. Epidemiol.* **157**, 98–112 (2003).
39. Sanderson, E. *et al.* Mendelian randomization. *Nat. Rev. Methods Primer* **2**, (2022).
40. Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A. & Yavitz, A. The rate of return to the HighScope Perry Preschool Program. J. *Public Econ.* **94**, 114–128 (2010).
41. Conti, G., Heckman, J. J. & Pinto, R. The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *Econ. J.* **126**, F28–F65 (2016).
42. Gutman, L. M. & Schoon, I. The impact of non-cognitive skills on outcomes for young people. *Educ. Endow. Found.* **59**, 2019 (2013).
43. Garcia, E. *The Need to Address Noncognitive Skills in the Education Policy Agenda.* https://www.epi.org/publication/the-need-to-address-noncognitive-skills-in-the-education-policy-agenda/ (2014).
44. Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B. & Borghans, L. Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success." OECD Education Working Papers, No. 110, OECD Publishing, Paris. (2014).
45. Moffitt, T. E. *et al.* A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci.* **108**, 2693–2698 (2011).
46. Heckman, J. J. Skill Formation and the Economics of Investing in Disadvantaged Children. *Soc. Sci.* **312**, 2005–2007 (2006).
47. Smithers, L. G. *et al.* A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes. *Nat. Hum. Behav.* **2**, 867–880 (2018).
48. Tucker-Drob, E. M., Briley, D. A., Engelhardt, L. E., Mann, F. D. & Harden, K. P. Genetically-mediated associations between measures of childhood character and academic achievement. *J. Pers. Soc. Psychol.* **111**, 790–815 (2016).
49. Kovas, Y. *et al.* Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personal. Individ. Differ.* **80**, 51–63 (2015).
50. Loehlin, J. C. *Genes and environment in personality development.* (Sage Publications, 1992).
51. Tucker-Drob, E. M. & Harden, K. P. Learning motivation mediates gene-by-socioeconomic status interaction on mathematics achievement in early childhood. *Learn. Individ. Differ.* **22**,

37–45 (2012).

52. Malanchini, M., Engelhardt, L. E., Grotzinger, A. D., Harden, K. P. & Tucker-Drob, E. M. 'Same But Different': Associations Between Multiple Aspects of Self-Regulation, Cognition, and Academic Abilities. *J. Pers. Soc. Psychol.* (2018) doi:10.1037/pspp0000224.
53. Liu, J. Z., Erlich, Y. & Pickrell, J. K. Case–control association mapping by proxy using family history of disease. *Nat. Genet.* **49**, 325–331 (2017).
54. Bowles, S. & Gintis, H. *Schooling In Capitalist America: Educational Reform And The Contradictions Of Economic Life.* (Basic Books, 1977).
55. Heckman, J. J. & Rubinstein, Y. The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *Am. Econ. Rev.* **91**, 145–149 (2001).
56. Ackerman, P. L., Kanfer, R. & Goff, M. Cognitive and Noncognitive Determinants and Consequences of Complex Skill Acquisition. J. *Exp. Psychol. Appl.* **1**, 270–304 (1995).
57. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513 (2019).
58. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
59. Belsky, D. W. & Harden, K. P. Phenotypic Annotation: Using Polygenic Scores to Translate Discoveries From Genome-Wide Association Studies From the Top Down. *Curr. Dir. Psychol. Sci.* **28**, 82–90 (2019).
60. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
61. Ligthart, L. *et al.* The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. *Twin Res. Hum. Genet.* 1–14 (2019) doi:10.1017/thg.2019.93.
62. Harris, K. M. *et al.* Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). Int. J. *Epidemiol.* **48**, 1415–1415k (2019).
63. Poulton, R., Moffitt, T. E. & Silva, P. A. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc. Psychiatry Psychiatr. Epidemiol.* **50**, 679–693 (2015).
64. Moffitt, T. E. & E-risk Team. Teen-aged mothers in contemporary Britain. *J. Child Psychol. Psychiatry* **43**, 727–742 (2002).
65. Herd, P., Carr, D. & Roan, C. Cohort profile: Wisconsin longitudinal study (WLS). *Int. J. Epidemiol.* **43**, 34–41 (2014).
66. Harden, K. P., Tucker-Drob, E. M. & Tackett, J. L. The Texas Twin Project. *Twin Res. Hum. Genet.* **16**, 385–390 (2013).
67. Benyamin, B. *et al.* Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol. Psychiatry* **19**, 253–258 (2014).
68. Chetty, R. *et al.* The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA* **315**, 1750 (2016).

69. Case, A. & Deaton, A. Mortality and Morbidity in the 21st Century. *Brook. Pap. Econ. Act.* **2017**, 397–476 (2017).
70. Hill, W. D. *et al.* Molecular Genetic Contributions to Social Deprivation and Household Income in UK Biobank. *Curr. Biol.* **26**, 3083–3089 (2016).
71. Timmers, P. R. *et al.* Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances. *eLife* **8**, 1–40 (2019).
72. Almlund, M., Duckworth, A. L., Heckman, J. & Kautz, T. Personality Psychology and Economics. in *Handbook of the Economics of Education* vol. 4 1–181 (Elsevier, 2011).
73. Borghans, L., Duckworth, A. L., Heckman, J. J. & Weel, B. ter. The Economics and Psychology of Personality Traits. J. *Hum. Resour.* **43**, 972–1059 (2008).
74. Rabin, M. A perspective on psychology and economics. *Eur. Econ. Rev.* 29 (2002).
75. Becker, A., Deckers, T., Dohmen, T., Falk, A. & Kosse, F. The Relationship Between Economic Preferences and Psychological Personality Measures. *Annu. Rev. Econ.* **4**, 453–478 (2012).
76. Linnér, R. K. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* **51**, 245–257 (2019).
77. Sanchez-Roige, S. *et al.* Genome-wide association study of delay discounting in 23,217 adult research participants of European ancestry. *Nat. Neurosci.* **21**, 16–18 (2018).
78. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in 700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
79. Furberg, H. *et al.* Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
80. Walters, R. K. *et al.* Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* **21**, 1656–1669 (2018).
81. Schumann, G. *et al.* KLB is associated with alcohol drinking, and its gene product β-Klotho is necessary for FGF21 regulation of alcohol preference. *Proc. Natl. Acad. Sci.* **113**, 14372–14377 (2016).
82. Pasman, J. A. *et al.* GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. *Nat. Neurosci.* **21**, 1161–1170 (2018).
83. Karlsson Linnér, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat. Neurosci.* **24**, 1367–1376 (2021).
84. Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* **48**, 1462–1472 (2016).
85. Lo, M.-T. *et al.* Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat. Genet.* **49**, 152–156 (2017).
86. John, O. P., Naumann, L. P. & Soto, C. J. Paradigm shift to the integrative Big Five Trait taxonomy. *Handb. Personal. Theory Res.* 114–158 (2008) doi:10.1016/S0191-8869(97)81000-8.

87. de Moor, M. H. M. *et al.* Meta-analysis of genome-wide association studies for personality. *Mol. Psychiatry* **17**, 337–349 (2012).
88. Caspi, A., Roberts, B. W. & Shiner, R. L. Personality development: stability and change. *Annu. Rev. Psychol.* **56**, 453–484 (2005).
89. Kessler, R. C. *et al.* Social consequences of psychiatric disorders, I: educational attainment. *Am. J. Psychiatry* 1026–1032 (1995).
90. Breslau, J., Lane, M., Sampson, N. & Kessler, R. C. Mental Disorders and Subsequent Educational Attainment in a US National Sample. *J. Psychiatr. Res*. **42**, 708–716 (2008).
91. Power, R. A. *et al.* Polygenic risk scores for schizophrenia and bipolar disorder predict creativity. *Nat. Neurosci.* **18**, 953–955 (2015).
92. Bansal, V. *et al.* Genome-wide association study results for educational attainment aid in identifying genetic heterogeneity of schizophrenia. *Nat. Commun. Lond.* **9**, 1–12 (2018).
93. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
94. Ruderfer, D. M. *et al.* Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* **173**, 1705-1715.e16 (2018).
95. Jansen, P. R. *et al.* Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).
96. Duncan, L. *et al.* Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am. J. Psychiatry* **174**, 850–858 (2017).
97. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
98. Arnold, P. D. *et al.* Revealing the complex genetic architecture of obsessive–compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
99. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
100. Bulik-Sullivan, B. *et al.* An Atlas of Genetic Correlations across Human Diseases and Traits. *Nat. Genet.* **47**, 1236–1241 (2015).
101. Nieuwboer, H. A., Pool, R., Dolan, C. V., Boomsma, D. I. & Nivard, M. G. GWIS: Genome-Wide Inferred Statistics for Functions of Multiple Phenotypes. *Am. J. Hum. Genet.* **99**, 917–927 (2016).
102. Ardlie, K. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
103. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, (2015).
104. Fehrmann, R. S. N. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47**, 115–125 (2015).
105. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set

Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).

106. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
107. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999-1014.e22 (2018).
108. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
109. Nave, G., Jung, W. H., Karlsson Linnér, R., Kable, J. W. & Koellinger, P. D. Are Bigger Brains Smarter? Evidence From a Large-Scale Preregistered Study. *Psychol. Sci.* **30**, 43–54 (2019).
110. Elliott, M. L. *et al.* A Polygenic Score for Higher Educational Attainment is Associated with Larger Brains. *Cereb. Cortex* **29**, 3496–3504 (2019).
111. Zhao, B. *et al.* Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nat. Genet.* **51**, 1637–1644 (2019).
112. Klein, A. & Tourville, J. 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. *Front. Neurosci.* **6**, 171 (2012).
113. Klein, A. Mindboggle-101 manually labeled individual brains. (2016) doi:10.7910/DVN/HMQKCK.
114. Gorgolewski, K. J. *et al.* NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinformatics* **9**, 8 (2015).
115. Zhao, B. *et al.* Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n = 17,706). *Mol. Psychiatry* (2019) doi:10.1038/s41380-019-0569-z.
116. Haushofer, J. & Fehr, E. On the psychology of poverty. *Science* **344**, 862–867 (2014).
117. Briley, D. A., Domiteaux, M. & Tucker-Drob, E. M. Achievement-relevant personality: Relations with the Big Five and validation of an efficient instrument. *Learn. Individ. Differ.* **32**, 26–39 (2014).
118. von Stumm, S., Hell, B. & Chamorro-Premuzic, T. The Hungry Mind: Intellectual Curiosity Is the Third Pillar of Academic Performance. *Perspect. Psychol. Sci.* **6**, 574–588 (2011).
119. Smoller, J. W. *et al.* Psychiatric genetics and the structure of psychopathology. *Mol. Psychiatry* **24**, 409–420 (2019).
120. Plomin, R., Haworth, C. M. A. & Davis, O. S. P. Common disorders are quantitative traits. *Nat. Rev. Genet.* **10**, 872–878 (2009).
121. Meehl, P. E. Schizotaxia, schizotypy, schizophrenia. *Am. Psychol.* **17**, 827–838 (1962).
122. von Stumm, S. & Ackerman, P. L. Investment and intellect: A review and meta-analysis. *Psychol. Bull.* **139**, 841–869 (2013).
123. Tucker-Drob, E. M. & Harden, K. P. A Behavioral Genetic Perspective on Non-Cognitive Factors and Academic Achievement. in *Genetics, Ethics and Education* 134–158 (Cambridge University Press, 2017). doi:10.1017/9781316340301.007.

124. Tucker-Drob, E. M. Motivational factors as mechanisms of gene-environment transactions in cognitive development and academic achievement. in *Handbook of competence and motivation: Theory and application, 2nd ed.* 471–486 (The Guilford Press, 2017).
125. Tucker-Drob, E. M. & Harden, K. P. Intellectual Interest Mediates Gene × Socioeconomic Status Interaction on Adolescent Academic Achievement: Intellectual Interest and G×E. *Child Dev.* no-no (2012) doi:10.1111/j.1467-8624.2011.01721.x.
126. Malanchini, M. *et al.* Reading self-perceived ability, enjoyment and achievement: A genetically informative study of their reciprocal links over time. *Dev. Psychol.* **53**, 698–712 (2017).
127. Westfall, J. & Yarkoni, T. Statistically Controlling for Confounding Constructs Is Harder than You Think. *PLOS ONE* **11**, e0152719 (2016).
128. de la Fuente, J., Davies, G., Grotzinger, A. D., Tucker-Drob, E. M. & Deary, I. J. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. *Nat. Hum. Behav.* (2020) doi:10.1038/s41562-020-00936-2.
129. Tucker-Drob, E. M. & Briley, D. A. Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychol. Bull.* **140**, 949–979 (2014).
130. Tropf, F. C. *et al.* Hidden heritability due to heterogeneity across seven populations. Nat. *Hum. Behav.* **1**, 757–765 (2017).
131. Duncan, L. *et al.* Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, (2019).
132. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
133. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
134. Willemsen, G. *et al.* The Adult Netherlands Twin Register: Twenty-Five Years of Survey and Biological Data Collection. *Twin Res. Hum. Genet.* **16**, 271–281 (2013).
135. Highland, H. M., Avery, C. L., Duan, Q., Li, Y. & Harris, K. M. *Quality control analysis of Add Health GWAS data.* https://www.cpc.unc.edu/projects/addhealth/documentation/guides/AH_GWAS_QC.pdf (2018).
136. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
137. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).
138. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
139. Deary, I. J. *et al.* Intergenerational social mobility and mid-life status attainment: Influences of childhood intelligence, childhood social factors, and education. *Intelligence* **33**, 455–472 (2005).

140. Oreopoulos, P. & Salvanes, K. G. Priceless: The Nonpecuniary Benefits of Schooling. *J. Econ. Perspect.* **25**, 159–184 (2011).
141. Heckman, J. J., Stixrud, J. & Urzua, S. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *J. Labor Econ.* **24**, 411–482 (2006).
142. Leerkes, E. M., Bailes, L. G. & Augustine, M. E. The intergenerational transmission of emotion socialization. *Dev. Psychol.* **56**, 390–402 (2020).
143. Cunha, F. & Heckman, J. The Technology of Skill Formation. *Am. Econ. Rev.* **97**, 31–47 (2007).
144. Guhin, J., Calarco, J. M. & Miller-Idriss, C. Whatever Happened to Socialization? *Annu. Rev. Sociol.* **47**, 109–129 (2021).
145. McGue, M., Rustichini, A. & Iacono, W. G. Cognitive, Noncognitive, and Family Background Contributions to College Attainment: A Behavioral Genetic Perspective. *J. Pers.* **85**, 65–78 (2017).
146. Doren, C. & Grodsky, E. What Skills Can Buy: Transmission of advantage through cognitive and noncognitive skills. *Sociol. Educ.* **89**, 321–342 (2016).
147. Grönqvist, E., Öckert, B. & Vlachos, J. The Intergenerational Transmission of Cognitive and Noncognitive Abilities. *J. Hum. Resour.* **52**, 887–918 (2017).
148. Crawford, C., Goodman, A. & Joyce, R. Explaining the socio-economic gradient in child outcomes: the inter-generational transmission of cognitive skills. *Longitud. Life Course Stud.* **2**, 77–93 (2011).
149. Gutman, L. M. & Schoon, I. *The impact of non-cognitive skills on outcomes for young people. A literature review. Education Endowment Foundation: London, UK.* https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/essential-life-skills/ (2013).
150. Tucker-Drob, E. M., Briley, D. A., Engelhardt, L. E., Mann, F. D. & Harden, K. P. Genetically-mediated associations between measures of childhood character and academic achievement. J. *Pers. Soc. Psychol.* **111**, 790–815 (2016).
151. Kovas, Y. *et al.* Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personal. Individ. Differ.* **80**, 51–63 (2015).
152. Demange, P. A. *et al.* Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
153. Zhang, G. *et al.* Assessing the Causal Relationship of Maternal Height on Birth Size and Gestational Age at Birth: A Mendelian Randomization Analysis. *PLoS Med.* **12**, e1001865 (2015).
154. Lawlor, D. *et al.* Using Mendelian randomization to determine causal effects of maternal pregnancy (intrauterine) exposures on offspring outcomes: Sources of bias and methods for assessing them. *Wellcome Open Res.* **2**, 11 (2017).
155. Bates, T. C. *et al.* The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families. *Twin Res. Hum. Genet.* **21**, 73–83 (2018).

156. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428 (2018).
157. de Zeeuw, E. L. *et al.* Intergenerational Transmission of Education and ADHD: Effects of Parental Genotypes. *Behav. Genet.* **50**, 221–232 (2020).
158. Selzam, S. *et al.* Comparing within- and between-family polygenic score prediction. *Am. J. Hum. Genet. 2019* 605006 (2019) doi:10.1101/605006.
159. Cheesman, R. *et al.* Comparison of Adopted and Nonadopted Individuals Reveals Gene–Environment Interplay for Education in the UK Biobank. *Psychol. Sci.* **31**, 582–591 (2020).
160. Fulker, D. W., Cherny, S. S., Sham, P. C. & Hewitt, J. K. Combined linkage and association sib-pair analysis for quantitative traits. *Am. J. Hum. Genet.* **64**, 259–267 (1999).
161. Abecasis, G. R., Cardon, L. R. & Cookson, W. O. C. A General Test of Association for Quantitative Traits in Nuclear Families. *Am. J. Hum. Genet.* **66**, 279–292 (2000).
162. Eaves, L. J., Pourcain, B. S. t., Smith, G. D. avey, York, T. P. & Evans, D. M. Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis ('M-GCTA'). *Behav. Genet.* **44**, 445–455 (2014).
163. Eilertsen, E. M. *et al.* Direct and Indirect Effects of Maternal, Paternal, and Offspring Genotypes: Trio-GCTA. *Behav. Genet.* **51**, 154–161 (2021).
164. Warrington, N. M., Freathy, R. M., Neale, M. C. & Evans, D. M. Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank. *Int. J. Epidemiol.* **47**, 1229–1241 (2018).
165. Warrington, N. M. *et al.* Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat. Genet.* **51**, 804–814 (2019).
166. Brumpton, B. *et al.* Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **11**, 3519 (2020).
167. Davies, N. M. *et al.* Within family Mendelian randomization studies. *Hum. Mol. Genet.* **28**, R170–R179 (2019).
168. Morris, T. T., Davies, N. M., Hemani, G. & Smith, G. D. Population phenomena inflate genetic associations of complex social traits. *Sci. Adv.* **6**, eaay0328 (2020).
169. Demange, P. & Cheesman, R. PerlineDemange/GeneticNurtureNonCog: (2022) doi:10.5281/zenodo.6581326.
170. Wang, B. *et al.* Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts. *Am. J. Hum. Genet.* **108**, 1780–1791 (2021).
171. Isungset, M. A. *et al.* Social and Genetic Effects on Educational Performance in Early Adolescence. Working Paper at https://doi.org/10.3386/w28498 (2021).
172. Tubbs, J. D., Porsch, R. M., Cherny, S. S. & Sham, P. C. The Genes We Inherit and Those We Don't: Maternal Genetic Nurture and Child BMI Trajectories. *Behav. Genet.* **50**, 310–319 (2020).
173. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).

174. Young, A. I. Solving the missing heritability problem. *PLoS Genet.* **15**, e1008222 (2019).
175. Kovas, Y. *et al.* Literacy and Numeracy Are More Heritable Than Intelligence in Primary School. *Psychol. Sci.* **24**, 2048–2056 (2013).
176. Silventoinen, K. *et al.* Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Sci. Rep.* **10**, 12681 (2020).
177. Equity in Education : Breaking Down Barriers to Social Mobility | PISA | OECD iLibrary. https://www-oecd-ilibrary-org.vu-nl.idm.oclc.org/education/equity-in-education_9789264073234-en (2018).
178. Spijker, F. V., Houwen, K. V. D. & Gaalen, R. V. Invloed ouderlijk opleidingsniveau reikt tot ver in het voortgezet onderwijs. in (2017).
179. Young, A. I. *et al.* Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nat. Genet.* **54**, 897–905 (2022).
180. Kong, A., Benonisdottir, S. & Young, A. I. Family Analysis with Mendelian Imputations. 2020.07.02.185181 Preprint at https://doi.org/10.1101/2020.07.02.185181 (2020).
181. Yengo, L. *et al.* Imprint of assortative mating on the human genome. *Nat. Hum. Behav.* **2**, 948–954 (2018).
182. Balbona, J. V., Kim, Y. & Keller, M. C. Estimation of Parental Effects Using Polygenic Scores. *Behav. Genet.* (2021) doi:10.1007/s10519-020-10032-w.
183. Kim, Y., Balbona, J. V. & Keller, M. C. Bias and Precision of Parameter Estimates from Models Using Polygenic Scores to Estimate Environmental and Genetic Parental Influences. *Behav. Genet.* (2020) doi:10.1007/s10519-020-10033-9.
184. Howe, L. J., Evans, D. M., Hemani, G., Smith, G. D. & Davies, N. M. Evaluating indirect genetic effects of siblings using singletons. *PLOS Genet.* **18**, e1010247 (2022).
185. Fletcher, J. M., Wu, Y., Zhao, Z. & Lu, Q. The production of within-family inequality: Insights and implications of integrating genetic data. 2020.06.06.137778 Preprint at https://doi.org/10.1101/2020.06.06.137778 (2020).
186. Lyngstad, T. H., Ystrøm, E. & Zambrana, I. M. An Anatomy of Intergenerational Transmission: Learning from the educational attainments of Norwegian twins and their parents. (2017) doi:10.31235/osf.io/fby2t.
187. Armstrong-Carter, E. *et al.* The Earliest Origins of Genetic Nurture: The Prenatal Environment Mediates the Association Between Maternal Genetics and Child Development. *Psychol. Sci.* **31**, 781–791 (2020).
188. Trejo, S. Exploring Genetic Influences on Birth Weight. Preprint at https://doi.org/10.31235/osf.io/7j59q (2020).
189. McGue, M. *et al.* The Environments of Adopted and Non-adopted Youth: Evidence on Range Restriction From the Sibling Interaction and Behavior Study (SIBS). *Behav. Genet.* **37**, 449–462 (2007).
190. Hwang, L.-D., Moen, G.-H. & Evans, D. M. Using adopted individuals to partition indirect

maternal genetic effects into prenatal and postnatal effects on offspring phenotypes. *eLife* **11**, e73671 (2022).

191. Ho, H., Plomin, R. & DeFries, J. C. Selective placement in adoption. *Soc. Biol.* **26**, 1–6 (1979).
192. Pirastu, N. *et al.* Genetic analyses identify widespread sex-differential participation bias. *Nat. Genet.* **53**, 663–671 (2021).
193. Fletcher, J., Wu, Y., Li, T. & Lu, Q. Interpreting Polygenic Score Effects in Sibling Analysis. *bioRxiv* 2021.07.16.452740 (2021) doi:10.1101/2021.07.16.452740.
194. Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
195. Pearson, R. M. *et al.* Maternal perinatal mental health and offspring academic achievement at age 16: the mediating role of childhood executive function. *J. Child Psychol. Psychiatry* **57**, 491–501 (2016).
196. Willems, Y. E., Boesen, N., Li, J., Finkenauer, C. & Bartels, M. The heritability of self-control: A meta-analysis. *Neurosci. Biobehav. Rev.* **100**, 324–334 (2019).
197. van Bergen, E., van Zuijen, T., Bishop, D. & de Jong, P. F. Why Are Home Literacy Environment and Children's Reading Skills Associated? What Parental Skills Reveal. *Read. Res. Q.* **52**, 147–160 (2017).
198. Lareau, A. *Unequal Childhoods: Class, Race, and Family Life, With an Update a Decade Later.* (2011).
199. Cooper, K. Are Poor Parents Poor Parents? The Relationship between Poverty and Parenting among Mothers in the UK. *Sociology* **55**, 349–383 (2021).
200. Allen, N. E., Sudlow, C., Peakman, T., Collins, R., & UK Biobank. UK biobank data: come and get it. *Sci. Transl. Med.* **6**, 224ed4 (2014).
201. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
202. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
203. Rimfeld, K. *et al.* Twins Early Development Study: A Genetically Sensitive Investigation into Behavioral and Cognitive Development from Infancy to Emerging Adulthood. *Twin Res. Hum. Genet. Off. J. Int. Soc. Twin Stud.* **22**, 508–513 (2019).
204. Abdellaoui, A. *et al.* Educational attainment influences levels of homozygosity through migration and assortative mating. *PLoS ONE* **10**, 1–14 (2015).
205. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
206. Abdellaoui, A. *et al.* Population structure, migration, and diversifying selection in the Netherlands. *Eur. J. Hum. Genet.* **21**, 1277–1285 (2013).
207. Abraham, G., Qiu, Y. & Inouye, M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinforma. Oxf. Engl*. **33**, 2776–2778 (2017).

208. Selzam, S. *et al.* Evidence for gene-environment correlation in child feeding: Links between common genetic variation for BMI in children and parental feeding practices. *PLOS Genet.* **14**, e1007757 (2018).
209. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
210. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
211. Bulik-Sullivan, B. *et al.* LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
212. Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R. & Dwyer, T. Regression models for twin studies: a critical review. *Int. J. Epidemiol.* **34**, 1089–1099 (2005).
213. Plomin, R. Assortative Mating by Unwed Biological Parents of Adopted Children. **1977**, 449–450 (1977).
214. Boardman, J. D. & Fletcher, J. M. To cause or not to cause? That is the question, but identical twins might not have all of the answers. *Soc. Sci. Med.* **127**, 198–200 (2015).
215. Kohler, H. P., Behrman, J. R. & Schnittker, J. Social science methods for twins data: Integrating causality, endowments, and heritability. *Biodemography Soc. Biol.* **57**, 88–141 (2011).
216. OECD. *Health at a Glance: Europe 2018: State of Health in the EU Cycle.* (Organisation for Economic Co-operation and Development, 2018).
217. World Health Organization. *World health statistics 2022: monitoring health for the SDGs, sustainable development goals.* (World Health Organization, 2022).
218. Glahn, D. C., Bearden, C. E., Bowden, C. L. & Soares, J. C. Reduced educational attainment in bipolar disorder. *J. Affect. Disord.* **92**, 309–312 (2006).
219. Lorant, V. *et al.* Socioeconomic Inequalities in Depression: A Meta-Analysis. *Am. J. Epidemiol.* **157**, 98–112 (2003).
220. Nordmo, M. *et al.* The educational burden of disease: a cohort study. *Lancet Public Health* **7**, e549–e556 (2022).
221. Sjölander, A. & Zetterqvist, J. Confounders, Mediators, or Colliders: What Types of Shared Covariates Does a Sibling Comparison Design Control For? *Epidemiology* **28**, 540–547 (2017).
222. Frisell, T., Öberg, S., Kuja-Halkola, R. & Sjölander, A. Sibling Comparison Designs: Bias From Non-Shared Confounders and Measurement Error. *Epidemiology* **23**, 713–720 (2012).
223. Sanderson, E. *et al.* Mendelian randomization. *Nat. Rev. Methods Primer* **2**, 6 (2022).
224. Munafò, M., Davies, N. M. & Davey Smith, G. Can genetics reveal the causes and consequences of educational attainment? *J. R. Stat. Soc. Ser. A Stat. Soc.* **183**, 681–688 (2020).
225. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. *Genet. Epidemiol.* **37**, 658–665 (2013).
226. Bowden, J., Smith, G. D. & Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).

227. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
228. Hartwig, F. P., Smith, G. D. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **46**, 1985–1998 (2017).
229. Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
230. Howe, L. J., Tudball, M., Davey Smith, G. & Davies, N. M. Interpreting Mendelian-randomization estimates of the effects of categorical exposures such as disease status and educational attainment. *Int. J. Epidemiol.* dyab208 (2021) doi:10.1093/ije/dyab208.
231. McFarland, M. J. & Wagner, B. G. Does a college education reduce depressive symptoms in American young adults? *Soc. Sci. Med.* **146**, 75–84 (2015).
232. Fujiwara, T. & Kawachi, I. Is education causally related to better health? A twin fixed-effect study in the USA. *Int. J. Epidemiol.* **38**, 1310–1322 (2009).
233. Halpern-Manners, A., Schnabel, L., Hernandez, E. M., Silberg, J. L. & Eaves, L. J. The Relationship between Education and Mental Health: New Evidence from a Discordant Twin Study. *Soc. Forces* **95**, 107–131 (2016).
234. Davies, N. M. *et al.* Multivariable two-sample Mendelian randomization estimates of the effects of intelligence and education on health. *eLife* **8**, e43990 (2019).
235. Viinikainen, J. *et al.* Does education protect against depression? Evidence from the Young Finns Study using Mendelian randomization. *Prev. Med.* **115**, 134–139 (2018).
236. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
237. Harrison, S. *et al.* The causal effects of health conditions and risk factors on social and socioeconomic outcomes: Mendelian randomization in UK Biobank. *Int. J. Epidemiol.* **49**, 1661–1681 (2020).
238. Adams, C. D. A multivariable Mendelian randomization to appraise the pleiotropy between intelligence, education, and bipolar disorder in relation to schizophrenia. *Sci. Rep.* **10**, 6018 (2020).
239. Dardani, C. *et al.* Is genetic liability to ADHD and ASD causally linked to educational attainment? *Int. J. Epidemiol.* **50**, 2011–2023 (2021).
240. Yuan, S., Xiong, Y., Michaëlsson, M., Michaëlsson, K. & Larsson, S. C. Genetically predicted education attainment in relation to somatic and mental health. *Sci. Rep.* **11**, 4296 (2021).
241. Pérez-Vigil, A. *et al.* Association of Obsessive-Compulsive Disorder With Objective Indicators of Educational Attainment. *JAMA Psychiatry* **75**, 47–55 (2018).
242. Vilaplana-Pérez, A. *et al.* Assessment of Posttraumatic Stress Disorder and Educational Achievement in Sweden. *JAMA Netw. Open* **3**, e2028477 (2020).

243. Lawlor, D. A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* **45**, 1866–1886 (2016).
244. Graeber, D. Does more education protect against mental health problems? *Dtsch. Inst. Für Wirtsch. DIW Berl.* **DIW Roundup: Politik im Fokus**, (2017).
245. Crespo, L., López-Noval, B. & Mira, P. Compulsory schooling, education, depression and memory: New evidence from SHARELIFE. *Econ. Educ. Rev.* **43**, 36–46 (2014).
246. Davies, N. M., Dickson, M., Davey Smith, G., Windmeijer, F. & van den Berg, G. J. The Causal Effects of Education on Adult Health, Mortality and Income: Evidence from Mendelian Randomization and the Raising of the School Leaving Age. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.3390179 (2019).
247. Davies, N. M., Dickson, M., Davey Smith, G., van den Berg, G. J. & Windmeijer, F. The causal effects of education on health outcomes in the UK Biobank. *Nat. Hum. Behav.* **2**, 117–125 (2018).
248. Weijer, M. V. de, Demange, P., Pelt, D., Bartels, M. & Nivard, M. G. Disentangling potential causal effects of educational duration on well-being, and mental and physical health outcomes. Preprint at https://doi.org/10.31234/osf.io/4uqyg (2022).
249. Hofmann, S. & Mühlenweg, A. Learning intensity effects in students' mental and physical health – Evidence from a large scale natural experiment in Germany. *Econ. Educ. Rev.* **67**, 216–234 (2018).
250. Root, A. *et al.* Association of Relative Age in the School Year With Diagnosis of Intellectual Disability, Attention-Deficit/Hyperactivity Disorder, and Depression. *JAMA Pediatr.* **173**, 1068–1075 (2019).
251. Böckerman, P., Haapanen, M., Jepsen, C. & Roulet, A. School Tracking and Mental Health. *J. Hum. Cap.*
252. Michaëlsson, M. *et al.* The impact and causal directions for the associations between diagnosis of ADHD, socioeconomic status, and intelligence by use of a bi-directional two-sample Mendelian randomization design. *BMC Med.* **20**, 106 (2022).
253. Kossek, E. E. & Lautsch, B. A. Work–Life Flexibility for Whom? Occupational Status and Work–Life Inequality in Upper, Middle, and Lower Level Jobs. *Acad. Manag. Ann.* (2017) doi:10.5465/annals.2016.0059.
254. Tempelaar, W. M., Termorshuizen, F., MacCabe, J. H., Boks, M. P. M. & Kahn, R. S. Educational achievement in psychiatric patients and their siblings: a register-based study in 30 000 individuals in The Netherlands. *Psychol. Med.* **47**, 776–784 (2017).
255. Vreeker, A. *et al.* High educational performance is a distinctive feature of bipolar disorder: a study on cognition in bipolar disorder, schizophrenia patients, relatives and controls. *Psychol. Med.* **46**, 807–818 (2016).
256. Sletved, K. S. O., Ziersen, S. C., Andersen, P. K., Vinberg, M. & Kessing, L. V. Socio-economic functioning in patients with bipolar disorder and their unaffected siblings – results from a

nation-wide population-based longitudinal study. *Psychol. Med.* 1–8 (2021) doi:10.1017/S0033291721002026.

257. MacCabe, J. H. *et al.* Artistic creativity and risk for schizophrenia, bipolar disorder and unipolar depression: A Swedish population-based case-control study and sib-pair analysis. *Br. J. Psychiatry* **212**, 370–376 (2018).
258. Power, R. A. *et al.* Polygenic risk scores for schizophrenia and bipolar disorder predict creativity. *Nat. Neurosci.* **18**, 953–955 (2015).
259. Smith, D. J. *et al.* Childhood IQ and risk of bipolar disorder in adulthood: prospective birth cohort study. *BJPsych Open* **1**, 74–80 (2015).
260. MacCabe, J. H. *et al.* Excellent school performance at age 16 and risk of adult bipolar disorder: national cohort study. *Br. J. Psychiatry J. Ment. Sci.* **196**, 109–115 (2010).
261. Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* **2**, 1–11 (2019).
262. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
263. Alten, S. van, Domingue, B. W., Galama, T. & Marees, A. T. Reweighting the UK Biobank to reflect its underlying sampling population substantially reduces pervasive selection bias due to volunteering. 2022.05.16.22275048 Preprint at https://doi.org/10.1101/2022.05.16.22275048 (2022).
264. Alten, S. van, Domingue, B. W., Faul, J., Galama, T. & Marees, A. T. Correcting for volunteer bias in GWAS uncovers novel genetic variants and increases heritability estimates. 2022.11.10.22282137 Preprint at https://doi.org/10.1101/2022.11.10.22282137 (2022).
265. Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M. & Davey Smith, G. Collider scope: when selection bias can substantially influence observed associations. Int. J. Epidemiol. 47, 226–235 (2018).
266. Burgess, S., Swanson, S. A. & Labrecque, J. A. Are Mendelian randomization investigations immune from bias due to reverse causation? *Eur. J. Epidemiol.* **36**, 253–257 (2021).
267. Elm, E. von *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet* **370**, 1453–1457 (2007).
268. Skrivankova, V. W. *et al.* Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *BMJ* **375**, n2233 (2021).
269. Zeeuw, E. L. de *et al.* Safe Linkage of Cohort and Population-Based Register Data in a Genomewide Association Study on Health Care Expenditure. *Twin Res. Hum. Genet.* **24**, 103–109 (2021).
270. Burgess, S. *et al.* Guidelines for performing Mendelian randomization investigations. Preprint

at https://doi.org/10.12688/wellcomeopenres.15555.2 (2020).

271. Davies, N. M. *et al.* Within family Mendelian randomization studies. *Hum. Mol. Genet.* **28**, R170–R179 (2019).
272. Demange, P. A. *et al.* Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. *Nat. Commun.* **13**, 4801 (2022).
273. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
274. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
275. R Core Team. R: A Language and Environment for Statistical Computing. (2021).
276. Burgess, S. & Labrecque, J. A. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur. J. Epidemiol.* **33**, 947–952 (2018).
277. Greco M, F. D., Minelli, C., Sheehan, N. A. & Thompson, J. R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* **34**, 2926–2940 (2015).
278. Burgess, S. & Thompson, S. G. Bias in causal estimates from Mendelian randomization studies with weak instruments. *Stat. Med.* **30**, 1312–1323 (2011).
279. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample mendelian randomization analyses using MR-Egger regression: The role of the I 2 statistic. *Int. J. Epidemiol.* **45**, 1961–1974 (2016).
280. Abel, K. M. *et al.* Prevalence of maternal mental illness among children and adolescents in the UK between 2005 and 2017: a national retrospective cohort analysis. *Lancet Public Health* **4**, e291–e300 (2019).
281. Bassani, D. G., Padoin, C. V., Philipp, D. & Veldhuizen, S. Estimating the number of children exposed to parental psychiatric disorders through a national health survey. *Child Adolesc. Psychiatry Ment. Health* **3**, 6 (2009).
282. Maybery, D. *et al.* Prevalence of parental mental illness in Australian families. *Psychiatr. Bull.* **33**, 22–26 (2009).
283. Pierce, M. *et al.* Prevalence of parental mental illness and association with socioeconomic adversity among children in Sweden between 2006 and 2016: a population-based cohort study. *Lancet Public Health* **5**, e583–e591 (2020).
284. Heradstveit, O. *et al.* Parental Mental Illness as a Risk Factor for Adolescent Psychiatric Disorders: A Registry-Based Study of Specialized Child and Adolescent Health Services. *Child Youth Serv.* **0**, 1–24 (2021).
285. Jami, E. S., Hammerschlag, A. R., Bartels, M. & Middeldorp, C. M. Parental characteristics and offspring mental health and related outcomes: a systematic review of genetically informative literature. *Transl. Psychiatry* **11**, (2021).
286. Raitasalo, K., Holmila, M., Jääskeläinen, M. & Santalahti, P. The effect of the severity of parental

alcohol abuse on mental and behavioural disorders in children. *Eur. Child Adolesc. Psychiatry* **28**, 913–922 (2019).

287. Ayano, G. *et al.* The impact of parental mental health problems on the educational outcomes of their offspring: Findings from the Raine Study. *Aust. N. Z. J. Psychiatry* **56**, 510–524 (2022).
288. Evans, J. *et al.* The timing of maternal depressive symptoms and child cognitive development: a longitudinal study. *J. Child Psychol. Psychiatry* **53**, 632–640 (2012).
289. Paget, A. *et al.* Which children and young people are excluded from school? Findings from a large British birth cohort study, the Avon Longitudinal Study of Parents and Children (ALSPAC). *Child Care Health Dev.* **44**, 285–296 (2018).
290. Berg, L., Charboti, S., Montgomery, E. & Hjern, A. Parental PTSD and school performance in 16-year-olds – a Swedish national cohort study. *Nord. J. Psychiatry* **73**, 264–272 (2019).
291. Brophy, S., Todd, C., Rahman, M. A., Kennedy, N. & Rice, F. Timing of parental depression on risk of child depression and poor educational outcomes: A population based routine data cohort study from Born in Wales, UK. *PLOS ONE* **16**, e0258966 (2021).
292. Lin, A. *et al.* Academic Performance in Children of Mothers With Schizophrenia and Other Severe Mental Illness, and Risk for Subsequent Development of Psychosis: A Population-Based Study. *Schizophr. Bull.* **43**, 205–213 (2017).
293. Netsi, E. *et al.* Association of Persistent and Severe Postnatal Depression With Child Outcomes. *JAMA Psychiatry* **75**, 247–253 (2018).
294. Shen, H. *et al.* Associations of Parental Depression With Child School Performance at Age 16 Years in Sweden. *JAMA Psychiatry* **73**, 239–246 (2016).
295. Ranning, A. *et al.* School performance from primary education in the adolescent offspring of parents with schizophrenia and bipolar disorder– a national, register-based study. *Psychol. Med.* **48**, 1993–2000 (2018).
296. Martini, M. G., Barona-Martinez, M. & Micali, N. Eating disorders mothers and their children: a systematic review of the literature. *Arch. Womens Ment. Health* **23**, 449–467 (2020).
297. Lees, B. *et al.* Association of Prenatal Alcohol Exposure With Psychological, Behavioral, and Neurodevelopmental Outcomes in Children From the Adolescent Brain Cognitive Development Study. *Am. J. Psychiatry* **177**, 1060–1072 (2020).
298. Zhang, H., Lee, Z. X., White, T. & Qiu, A. Parental and social factors in relation to child psychopathology, behavior, and cognitive function. *Transl. Psychiatry* **10**, 1–9 (2020).
299. Demange, P. A., Boomsma, D. I., Bergen, E. van & Nivard, M. G. Evaluating the causal relationship between educational attainment and mental health. 2023.01.26.23285029 Preprint at https://doi.org/10.1101/2023.01.26.23285029 (2023).
300. Kuja-Halkola, R., D'Onofrio, B. M., Larsson, H. & Lichtenstein, P. Maternal smoking during pregnancy and adverse outcomes in offspring: genetic and environmental sources of covariance. *Behav. Genet.* **44**, 456–467 (2014).
301. Jundong, J. *et al.* Poor school performance in offspring of patients with schizophrenia: what

are the mechanisms? *Psychol. Med.* **42**, 111–123 (2012).
302. phenotools. https://github.com/psychgen/phenotools (2023).
303. Derogatis, L. R., Lipman, R. S. & Covi, L. SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacol. Bull.* **9**, 13–28 (1973).
304. Tambs, K. & Røysamb, E. Selection of questions to short-form versions of original psychometric instruments in MoBa. *Nor. Epidemiol.* **24**, (2014).
305. Kessler, R. C. *et al.* The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol. Med.* **35**, 245–256 (2005).
306. Kessler, R. C. *et al.* Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) Screener in a representative sample of health plan members. *Int. J. Methods Psychiatr. Res.* **16**, 52–65 (2007).
307. DeRoo, L. A. Maternal alcohol consumption in the Norwegian Mother and Child Cohort Study (MoBa) – Research opportunities. *Nor. Epidemiol.* **24**, (2014).
308. Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R. & Grant, M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addict. Abingdon Engl.* **88**, 791–804 (1993).
309. *Diagnostic and statistical manual of mental disorders, 4th ed.* xxvii, 886 (American Psychiatric Publishing, Inc., 1994).
310. Bulik, C. M. *et al.* Patterns of remission, continuation and incidence of broadly defined eating disorders during early pregnancy in the Norwegian Mother and Child Cohort Study (MoBa). *Psychol. Med.* **37**, 1109–1118 (2007).
311. Fekadu, W., Mihiretu, A., Craig, T. K. J. & Fekadu, A. Multidimensional impact of severe mental illness on family members: systematic review. *BMJ Open* **9**, e032391 (2019).
312. Ranning, A., Munk Laursen, T., Thorup, A., Hjorthøj, C. & Nordentoft, M. Children of Parents With Serious Mental Illness: With Whom Do They Grow Up? A Prospective, Population-Based Study. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 953–961 (2016).
313. Ayano, G., Betts, K., Dachew, B. A. & Alati, R. Academic performance in adolescent offspring of mothers with prenatal and perinatal psychiatric hospitalizations: A register-based, data linkage, cohort study. *Psychiatry Res.* **319**, 114946 (2023).
314. Kleppesto, T. H. *et al.* Intergenerational transmission of ADHD behaviors: genetic and environmental pathways. Preprint at https://doi.org/10.31234/osf.io/ng5b8 (2023).
315. Pingault, J.-B. *et al.* Genetic nurture versus genetic transmission of risk for ADHD traits in the Norwegian Mother, Father and Child Cohort Study. *Mol. Psychiatry* 1–8 (2022) doi:10.1038/s41380-022-01863-6.
316. Nordmo, M. *et al.* The educational burden of disease: a cohort study. *Lancet Public Health* **7**, e549–e556 (2022).
317. Eley, T. C. *et al.* The Intergenerational Transmission of Anxiety: A Children-of-Twins Study. *Am.*

J. Psychiatry **172**, 630–637 (2015).

318. Ask, H. *et al.* Intergenerational transmission of parental neuroticism to emotional problems in 8-year-old children: Genetic and environmental influences. *JCPP Adv.* **1**, e12054 (2021).
319. Kennewell, E., Curtis, R. G., Maher, C., Luddy, S. & Virgara, R. The relationships between school children's wellbeing, socio-economic disadvantage and after-school activities: a cross-sectional study. *BMC Pediatr.* **22**, 297 (2022).
320. Calling, S., Ohlsson, H., Sundquist, J., Sundquist, K. & Kendler, K. S. Socioeconomic status and alcohol use disorders across the lifespan: A co-relative control study. *PLOS ONE* **14**, e0224127 (2019).
321. Huang, B., Shawn Chang, Y.-H., Niu, L. & Zhi, M. Examining the effects of socio-economic status and language input on adolescent English learners' speech production outcomes. *System* **73**, 27–36 (2018).
322. Flak, A. L. *et al.* The Association of Mild, Moderate, and Binge Prenatal Alcohol Exposure and Child Neuropsychological Outcomes: A Meta-Analysis. *Alcohol Clin. Exp. Res.* **38**, 214–226 (2014).
323. Bloomfield, K. Understanding the alcohol-harm paradox: what next? *Lancet Public Health* **5**, e300–e301 (2020).
324. Torvik, F. A., Rognmo, K. & Tambs, K. Alcohol use and mental distress as predictors of non-response in a general population health survey: the HUNT study. Soc. Psychiatry Psychiatr. Epidemiol. 47, 805–816 (2012).
325. Kinge, J. M. *et al.* Parental income and mental disorders in children and adolescents: prospective register-based study. *Int. J. Epidemiol.* **50**, 1615–1627 (2021).
326. Frisell, T., Öberg, S., Kuja-Halkola, R. & Sjölander, A. Sibling Comparison Designs: Bias From Non-Shared Confounders and Measurement Error. *Epidemiology* **23**, 713–720 (2012).
327. Lund, I. O. *et al.* Maternal Drinking and Child Emotional and Behavior Problems. *Pediatrics* **145**, (2020).
328. Sjölander, A., Frisell, T. & Öberg, S. Sibling Comparison Studies. *Annu. Rev. Stat. Its Appl.* **9**, 71–94 (2022).
329. Esping-Andersen, Gøsta. The Three Worlds of Welfare Capitalism. (1990).
330. Caspi, A., Roberts, B. W. & Shiner, R. L. Personality Development: Stability and Change. *Annu. Rev. Psychol.* **56**, 453–484 (2005).
331. Dohmen, T., Falk, A., Golsteyn, B. H. H., Huffman, D. & Sunde, U. Risk Attitudes across the Life Course. *Econ. J.* **127**, F95–F116 (2017).
332. Kureishi, W., Paule-Paludkiewicz, H., Tsujiyama, H. & Wakabayashi, M. Time Preferences over the Life Cycle. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.3547301 (2020).
333. Nordmo, M. *et al.* The Association Between Parental Internalizing Disorders and Child School Performance. *Prep.*
334. Weye, N. *et al.* Agreement between survey- and register-based measures of depression in

Denmark. *Acta Psychiatr. Scand.* **n/a,**.

335. Breznau, N. *et al.* Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proc. Natl. Acad. Sci.* **119**, e2203150119 (2022).
336. Ramus, F. *Noncognitive nonsense.* https://osf.io/xzv4f (2022) doi:10.31234/osf.io/xzv4f.
337. Mitchell, B. L. *et al.* Polygenic influences associated with adolescent cognitive skills. *Intelligence* **94**, 101680 (2022).
338. Eising, E. *et al.* Genome-wide analyses of individual differences in quantitatively assessed reading- and language-related skills in up to 34,000 people. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2202764119 (2022).
339. Thorp, J. G. *et al.* Genetic evidence that the causal association of educational attainment with reduced risk of Alzheimer's disease is driven by intelligence. *Neurobiol. Aging* **119**, 127–135 (2022).
340. Lam, M. *et al.* Collective genomic segments with differential pleiotropic patterns between cognitive dimensions and psychopathology. *Nat. Commun.* **13**, 6868 (2022).
341. Malanchini, M. *et al.* Genetic contributions of noncognitive skills to academic development. 2023.04.03.535380 Preprint at https://doi.org/10.1101/2023.04.03.535380 (2023).
342. Abdellaoui, A., Dolan, C. V., Verweij, K. J. H. & Nivard, M. G. Gene–environment correlations across geographic regions affect genome-wide association studies. *Nat. Genet.* **54**, 1345–1354 (2022).
343. Le, H. T. & Nguyen, H. T. Parental health and children's cognitive and noncognitive development: New evidence from the longitudinal survey of Australian children. *Health Econ.* **26**, 1767–1788 (2017).
344. Blanden, J., Gregg, P. & Macmillan, L. Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *Econ. J.* **117**, C43–C60 (2007).
345. Del Bono, E., Kinsler, J. & Pavan, R. *Skill Formation and the Trouble with Child Non-Cognitive Skill Measures.* https://www.econstor.eu/handle/10419/227240 (2020).
346. Jami, E. S. *et al.* Genome-wide Association Meta-analysis of Childhood and Adolescent Internalizing Symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* **61**, 934–945 (2022).
347. Akingbuwa, W. A. *et al.* Genetic Associations Between Childhood Psychopathology and Adult Depression and Associated Traits in 42 998 Individuals: A Meta-Analysis. *JAMA Psychiatry* (2020) doi:10.1001/jamapsychiatry.2020.0527.
348. Goodman, R. Strengths and Difficulties Questionnaire. (1997) doi:10.1037/t00540-000.
349. Achenbach, T. M. The Child Behavior Checklist and related instruments. in *The use of psychological testing for treatment planning and outcomes assessment, 2nd ed* 429–466 (Lawrence Erlbaum Associates Publishers, 1999).
350. Caro, J. C. *et al.* Children's internalizing behavior development is heterogeneously associated with the pace of epigenetic aging. *Biol. Psychol.* **176**, 108463 (2023).
351. Menta, G., Lepinteur, A., Clark, A. E., Ghislandi, S. & D'Ambrosio, C. Maternal genetic risk for

depression and child human capital. *J. Health Econ.* **87**, 102718 (2023).

352. Smit, D. J. A. *et al.* A genome-wide association study of a rage-related misophonia symptom and the genetic link with audiological traits, psychiatric disorders, and personality. *Front. Neurosci.* **16**, (2023).
353. Ahangari, M. *et al.* Examining the source of increased bipolar disorder and major depressive disorder common risk variation burden in multiplex schizophrenia families. *Schizophr. Heidelb. Ger.* **8**, 106 (2022).
354. Zhang, F., Cao, H. & Baranova, A. Genetic variation mediating neuroticism's influence on cardiovascular diseases. *J. Psychopathol. Clin. Sci.* **131**, 278–286 (2022).
355. Vries, L. P. de *et al.* Distinguishing Happiness and Meaning in Life from Depressive Symptoms: a GWAS-by-subtraction study in the UK Biobank. 2022.12.06.519260 Preprint at https://doi.org/10.1101/2022.12.06.519260 (2022).
356. Fitzgerald, J. *et al.* Thirteen Independent Genetic Loci Associated with Preserved Processing Speed in a Study of Cognitive Resilience in 330,097 Individuals in the UK Biobank. *Genes* **13**, 122 (2022).
357. Nivard, M. G. *et al.* Neither nature nor nurture: Using extended pedigree data to elucidate the origins of indirect genetic effects on offspring educational outcomes. Preprint at https://doi.org/10.31234/osf.io/bhpm5 (2022).
358. Fletcher, J., Wu, Y., Li, T. & Lu, Q. Interpreting Polygenic Score Effects in Sibling Analysis. 2021.07.16.452740 Preprint at https://doi.org/10.1101/2021.07.16.452740 (2021).
359. Boomsma, D. I. Sibling Interaction Effects. in *Encyclopedia of statistics in behavioral science* (John Wiley & Sons, 2005).
360. Rietveld, M. J. H., Posthuma, D., Dolan, C. V. & Boomsma, D. I. ADHD: Sibling Interaction or Dominance: An Evaluation of Statistical Power. *Behav. Genet.* **33**, 247–255 (2003).
361. Isungset, M. A., Freese, J., Andreassen, O. A. & Lyngstad, T. H. Birth order differences in education originate in postnatal environments. *PNAS Nexus* **1**, pgac051 (2022).
362. Luningham, J. M. *et al.* Harmonizing behavioral outcomes across studies, raters, and countries: application to the genetic analysis of aggression in the ACTION Consortium. *J. Child Psychol. Psychiatry* **61**, 807–817 (2020).
363. Gil-Hernández, C. J. Do Well-off Families Compensate for Low Cognitive Ability? Evidence on Social Inequality in Early Schooling from a Twin Study. *Sociol. Educ.* **92**, 150–175 (2019).
364. Simonoff, E. *et al.* Genetic influences on childhood hyperactivity: contrast effects imply parental rating bias, not sibling interaction. *Psychol. Med.* **28**, 825–837 (1998).
365. Lahey, B. B. & D'Onofrio, B. M. All in the Family: Comparing Siblings to Test Causal Hypotheses Regarding Environmental Influences on Behavior. *Curr. Dir. Psychol. Sci.* **19**, 319–323 (2010).
366. Madole, J. W. & Harden, K. P. Building Causal Knowledge in Behavior Genetics. *Behav. Brain Sci.* 1–76 (2022) doi:10.1017/S0140525X22000681.
367. Sjölander, A., Frisell, T., Kuja-Halkola, R., Öberg, S. & Zetterqvist, J. Carryover Effects in Sibling

Comparison Designs. *Epidemiology* **27**, 852 (2016).

368. Veller, C. & Coop, G. Interpreting population and family-based genome-wide association studies in the presence of confounding. 2023.02.26.530052 Preprint at https://doi.org/10.1101/2023.02.26.530052 (2023).
369. Wang, Y., Tsuo, K., Kanai, M., Neale, B. M. & Martin, A. R. Challenges and Opportunities for Developing More Generalizable Polygenic Risk Scores.
370. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat. Hum. Behav.* (2021) doi:10.1038/s41562-021-01110-y.
371. Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* **2**, 9 (2019).
372. Martin, A. R. *et al.* Increasing diversity in genomics requires investment in equitable partnerships and capacity building. *Nat. Genet.* **54**, 740–745 (2022).
373. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
374. Fatumo, S. *et al.* A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
375. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).
376. Vink, J. M. *et al.* Estimating non-response bias in family studies: application to mental health and lifestyle. *Eur. J. Epidemiol.* **19**, 623–630 (2004).
377. Schoeler, T. *et al.* Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nat. Hum. Behav.* 1–12 (2023) doi:10.1038/s41562-023-01579-9.
378. Shaw, R. J. *et al.* Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. *Eur. J. Epidemiol.* **37**, 1215–1224 (2022).
379. Stronks, K. *et al.* Do conceptualisations of health differ across social strata? A concept mapping study among lay people. *BMJ Open* **8**, e020210 (2018).
380. About ODISSEI. ODISSEI - *Open Data Infrastructure for Social Science and Economic Innovations* https://odissei-data.nl/en/en-odissei/.
381. Tryggve - NeIC web. https://neic.no/tryggve/.

AUTHOR CONTRIBUTIONS & COMPETING INTERESTS

CHAPTER 2

D.W.B., K.P.H., M.G.N., P.A.D., and M.M. conceived the idea for the study with assistance from E.M.T.-D., B.W.D., P.B, C.M., and J.W. P.A.D., M.M., T.T.M., P.B., B.W.D., D.W.B., D.L.C., K.S., S.R.C., M.G.N., A.A., and H.F.I. analysed the data (P.A.D ran the GWASs, the genetic correlations for the phenotypic annotation, the PGS analyses in NTR, and the different enrichment analyses.). D.W.B., K.P.H., M.G.N., M.M., P.A.D., and E.M.T.-D. wrote the paper with helpful contributions from P.B., B.W.D., and S.R.C. A.D.G., L.A., E.v.B., D.I.B., A.C., K.M.H., T.E.M., R.P., J.A.P., B.S.W., E.L.Z. and previously mentioned authors contributed to interpretation of data, provided critical feedback on manuscript drafts and approved the final draft.

The authors declare no competing interests.

CHAPTER 3

R.C. & P.A.D. conceived and designed the study, with helpful contributions from M.G.N. P.A.D. & R.C. analysed the data, with help from J.J.H. to obtain polygenic score weights and A.A. to perform GWAS in UK Biobank. P.A.D., M.G.N., R.C., and E.M.E. performed the simulation study. R.C. & P.A.D. wrote the manuscript. J.J.H., A.A., E.M.E., M.M., B.W.D., E.A.C., E.L.d.Z., K.R., D.I.B., E.v.B., and G.B. contributed to the interpretation of data, provided feedback on manuscript drafts, and approved the final draft.

The authors declare no competing interests.

CHAPTER 4

P.A.D. and M.G.N. conceived and designed the study, with helpful suggestions from E.v.B. and D.I.B. P.A.D. analysed the data, with support from M.G.N. for the MR analyses. P.A.D. designed the figures and drafted the manuscript. All authors contributed to and approved the final version of the manuscript.

The authors declare no competing interests.

CHAPTER 5

P.A.D, M.G.N, T.L., F.T., E.v.B. and E.Y conceived the study. P.A.D analysed the data, with support from E.Y. for the IRT analyses. P.A.D. designed the figures and drafted the manuscript. All authors contributed and approved the final version of the manuscript.

The authors declare no competing interests.

LIST OF PUBLICATIONS

De Vries L.P., **Demange P.A.**, Baselmans B.M.L., Vinkers C.H., Pelt D.H.M., Bartels M. (2023). Distinguishing happiness and meaning in life from depressive symptoms: a GWAS-by-subtraction study in the UK Biobank. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. https://doi.org/10.1002/ajmg.b.32954

Skeldal, S., Voss, L.F., Lende, J., Pedersen, S.B., Mølgaard, S., Ollendorff, M.K., **Demange, P.A.**, Bentsen, A. H., Fuglsang, M., Sander, M.R., Buttenschøn, H., Gustafsen, C., Madsen, P., & Glerup, S. (2023). Alternative splicing regulates adaptor protein binding, trafficking, and activity of the Vps10p domain receptor SorCS2 in neuronal development. Journal of Biological Chemistry, 0(0). https://doi.org/10.1016/j.jbc.2023.105102

Demange, P.A., Hottenga, J.J., Abdellaoui, A., Eilertsen, E.M., Malanchini, M., Domingue, B.W., Armstrong-Carter, E., de Zeeuw, E.L., Rimfeld, K., Boomsma, D.I., van Bergen, E., Breen, G., Nivard, M.G., & Cheesman, R. (2022). Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-32003-x

Pasman, J.A., **Demange, P.A.**, Guloksuz, S., Willemsen A.H.M., Abdellaoui A., ten Have M., Hottenga J.J., Boomsma D.I., de Geus E., Bartels M., de Graaf R., Verweij K.J.H., Smit D.J., Nivard M., Vink J.M. (2021). Genetic Risk for Smoking: Disentangling Interplay Between Genes and Socioeconomic Status. Behavior Genetics https://doi.org/10.1007/s10519-021-10094-4

Demange, P.A.*, Malanchini, M.*, Mallard, T.T., Biroli, P., Cox, S.R., Grotzinger, A.D., Tucker-Drob, E.M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D.I., Caspi, A., Corcoran, D.L., Domingue, B.W., Harris, K.M., Ip, H.F., Mitchell, C., Moffitt, T.E., Poulton, R., Prinz, J.A., Sugden, K., Wertz, J., Williams, B.S., de Zeeuw, E.L., Belsky, D.W*., Harden, K.P.* & Nivard, M.G*. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. Nature Genetics, 53(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Mell, H., Safra, L., **Demange, P.A.**, Algan, Y., Baumard, N., & Chevallier, C. (2021). Early life adversity is associated with diminished social trust in adults. Political Psychology https://doi.org/10.1111/pops.12756

Lawn, R.B., Sallis, H.M., Wootton, R.E., Taylor, A.E., **Demange, P.A.**, Fraser, A., Penton-Voak, I.S., & Munafò, M.R. (2020). The effects of age at menarche and first sexual intercourse on reproductive and behavioural outcomes: A Mendelian randomization study. PLoS ONE, 15(6 June), 1–17. https://doi.org/10.1371/journal.pone.0234488

Demange, P.A., Boomsma, D.I., van Bergen, E*., Nivard, M.G.* (2023) Evaluating the causal relationship between educational attainment and mental health. medRxiv, 2023.01.26. https://doi.org/10.1101/2023.01.26.23285029 (under review)

Van de Weijer, M., **Demange, P.A.**, Pelt, D., Bartels, M., and Nivard, M.G. (2022). Disentangling potential causal effects of educational duration on well-being, and mental and physical health outcomes. PsyArXiv, September 27. https://doi.org/10.31234/osf.io/4uqyg (under review)

Malanchini M., Allegrini A.G., Nivard M.G., Biroli P., Rimfeld K., Cheesman R., von Stumm S., **Demange P.A.**, van Bergen E., Grotzinger A.D., Raffington L., De la Fuente J., Pingault J-B., K.P., Tucker-Drob E.M. & Plomin R. Genetic contributions of noncognitive skills to academic development. bioRxiv, 2023.04.05. https://doi.org/10.1101/2023.04.03.535380 (under review)

Nordmo M., Kleppestø T., Sunde H.F., Flatø M., **Demange P.A.**, Torvik F.A. The Association Between Parental Internalizing Disorders and Child School Performance. (accepted)

Demange, P.A., Nivard, M. G., Torvik, F., Eilertsen, E.M., Cheesman. R., Nordmo, M., Lyngstad, T., Ystrøm, E.*, van Bergen, E.* No effect of parental mental health on children's academic achievement: a within-family study. (in preparation)

Ghirardi, G., Gil-Hernández, C.J., Bernardi, F., van Bergen, E.*, **Demange, P.A.*** Do high-SES parents compensate for genetic risk for poor educational achievement in offspring? A pre-registered genetically-sensitive study. (in preparation)

ACKNOWLEDGEMENTS

During my interview for this job, Michel asked me: «What's your aim for your PhD?»
I confidently replied: «To become a better scientist.»
Michel burst into a loud and surprised laughter. Seeing my puzzlement, he quickly clarified that this was indeed the most important goal for a PhD.
At that moment, I knew I was in the right place for my PhD journey.

This chapter is dedicated to all the incredible individuals who have not only helped me become a better scientist but also a better person. I have learned so much in the last five years. You've shaped my path in ways I would never have imagined, and I am deeply grateful for the profound impact you've had on my life.

To my supervisors, your unwavering belief in my abilities has been instrumental in my development. I am deeply grateful for the opportunities and challenges you have provided me. Michel, I cannot express enough gratitude for your enthusiasm, patience, and trust throughout my PhD journey. Being your student has been a stroke of luck, and your commitment to constant improvement as a scientist, supervisor, and person has been truly inspiring. Your ability to listen, your sense of humour, and your honesty make every conversation a pleasure. You have created an environment where I felt supported and encouraged to explore and develop my own scientific path. From you, I am learning the value of remaining curious and open to learning, not just about the subject matter but about myself as well. Elsje, your meticulous attention to scientific communication has left a lasting impact on me. Your check-ins during the lockdowns have helped me feel connected during difficult times, and your exemplary model of perseverance, boundaries, and work-life balance serves as a guide for me to follow. Dorret, it has been an immense honour to work with you and to benefit from your wealth of knowledge. Your reminder to leverage decades of scientific wisdom has been invaluable in shaping my research approach. Thank you for always paying attention to everyone and everything while giving me the freedom I needed in this project.
Eveline, it is unfortunate I only got to experience working with you for a few years! I am glad you found a position that suits you better, reminding us that academia is not an end in itself.
To all the members of my reading committee: Danielle Posthuma, Hans van Kippersluis, Pearl Dykstra, Herman van de Werfhorst, Laura Wesseldijk, Niels van den Berg, I sincerely thank you for the time and effort you dedicated to reading and evaluating my thesis.
To my paranymphs, thank you for being by my side in these last moments of my PhD journey. Anne, from my first day, you have been the wise voice reminding me to be kind, be brave, and rest. Now years after you left the department, your impact is still visible: you taught us PhDs to open up to each other in order to better support each other. Coaching was made for

you, and I am so glad you found it. Nicole, thank you for sharing and easing my frustrations. Your energy and determination to make things better, without compromising, always make me feel like I can make it too.

I want to thank all my colleagues at the Biological Psychology Department and Netherlands Twin Register. In particular, I want to thank all the PhD candidates. Thank you for all the help, but more importantly thank you for all the laughs. Your camaraderie was invaluable (and sorry for having no self-control and talking so much; I hope I was entertaining if not efficient). I am so proud of you all.
My TTT girls, if the hardships of a PhD, of a pandemic, of societal upheavals, brought one good thing, it is you. Thank you for all the good times. It is hard to realize how much we've grown in this time. Thank you for being so accepting, for embracing (y)ourselves, for loving each other enough to call ourselves out when needed. Thank you for all the TikToks, quirks, fandoms, healing our adolescent selves. Wonu, I can't wait for you to be a PhD mama, no need to sweat baby. Sofieke, you don't have to be sorry for leaving and growing up. Zenab, almost there, do not ever lose your fire. Margot, you have the skin of a killer.
Denise, thank you for the ADHD chatters, and the great example of determination and work/life balance. Eshim, I am so proud I got to be your paranymph, I miss you. Anke, I hope I can be an inspiration to others as much as you were to me. Your path is insightful and your work for UU is an example for all. Yayouk, Bart, thank you for the mentorship. Matthijs, Hill, thank you for saving my code so many times. Floris, I already miss seeing you across my desk every day. Selim, Nikki, thanks for the coffee breaks. Zoey, thank you for your kindness and food. Lianne, your efficiency scared me, thank you for being humble and trying to show us the way. Fiona, Susanne, Fiona, Veronika, Hekmat, Camiel, Bodine, Melissa, Myrte, Sjors, thank you for all the good times. Britt, Ora, Nina, Emily, thank you for letting me help you during your master's projects.

Academia and science are nothing without a team. I am so grateful I had amazing collaborators throughout my PhD. Rosa, could you imagine we could have missed each other at BGA? I am so thankful we embraced serendipity. I am always learning so much from you. You are the perfect teammate and friend, I am looking forward to many more projects. Carlos, thank you for your trust in this random geneticist attending a sociology conference. And thank you for bringing Gaia into my life. Gaia, your vibrant presence and boundless positivity are filling me with so much energy and joy. I am learning so much from this interdisciplinary collaboration, thank you for your hard work and your patient communication. Margherita, I could not have dreamt of a better collaborator for my first project. Your optimism and openness made me

immediately feel at home. Paige and Dan, I learned so much about writing from you, it was an honour to work with you both. Ben, thank you for your unfaltering enthusiasm and support. Small gestures of attention can have a big impact, and you are so generous with them, as well as with your time and ideas. To the AUMC and economics colleagues, thank you for the fun collaborations and journal clubs. To all the PROMENTA and HUP people, thank you for the amazing welcome in 2021, I am so looking forward to now joining for a postdoc! I have learned so much during this research visit, I enjoyed so many insightful conversations, and got so much help from all of you. Eivind, thank you for your supervision, for your enthusiasm, and for your patience and flexibility as I completed my thesis. Torkild and Fartein, I am looking forward to continuing the collaboration with you and your teams.
One of my favourite things about the job is meeting all these bright curious international minds at conferences and visits. Thank you to everyone I met who ever so briefly shared with me and made me feel like I belong. I had a blast at every conference I attended, every invited talk I gave. A special acknowledgement to those striving for safer and more inclusive conferences – your efforts enable these positive experiences. Andrea, conferences are always the best when you are there; who else would brainstorm ideas, give me feedback, while also reminding me to have a drink (of water)? Thank you to the IEU crowd who adopted me for a few days at ECSR. Felix, thank you for the never-ending optimism for me and my career. Robbee, Sjoerd, thank you for the good times on our Midwest tour.

A university is so much more than researchers and lecturers. I was lucky to discover a bit of this hidden world and the amazing people who work to give us researchers the means to give our best. Thank you to all administrative staff, especially the IO PhD coordinators and student assistants, the VU PhD advisors, the graduate school employees, and IT for such a quick answer during the pandemic. Natascha and Michiel, thank you for always being there to answer questions and organize a Christmas party. Sandra, thank you for getting me into PhD Gatherings and pushing me to represent the VU at PNN and the graduate school network. Terence and Merel, thank you for making PhD Gatherings as fun as possible, and for your efforts to keep connections between PhD students even during the pandemic. Thank you to all PhD students who joined us in these social events and opened up our tiny academic world. Rosanne, thank you for giving us the nudge to get a PhD council started and for all the work that went into it. Margot, Timothy, Nina, thank you for bringing the council to life. To Nicole and Floris, thank you for making it thrive. To all the PhDs involved, thank you so much for your time, energy, emotional labour, and your withstanding motivation to make work-life better for all. To Finette, thank you for being a strong support to the council from the start. To my peer coaching group, Hannah, Anita, Tessa, thank you for your presence,

your vulnerability, your courage and your advice, our meetings have taught me so much. To Sander and Lidy, thank you for going beyond your job description to make this happen. To the therapists and coaches I have seen in the last years, including at the VU Henk, Victor, and Samula, thank you for your help in keeping me afloat, mending the cracks in my prow.

Thank you to all that brought me to start a PhD, keeping my motivation and curiosity alive through my education. Thank you to all my previous supervisors, mentors, and teachers. Sune, I still hear you telling me “Just give it a go” each time I am overthinking and being scared of starting a new project. Lars, thank you for offering me to be involved so many times; you make me feel like an independent researcher since my Bachelor. Coralie, thank you for the hours spent on my master’s thesis, your feedback on my presentations is still the fundamental layer of my communication skills. Marcus, thank you for introducing me to genetic epidemiology and reproducibility, and supporting me ever since. Thank you to all for paving the way that brought me to this specific PhD position, Ruben for pointing me towards Bristol, Jorien for spotting this Dutch job advertisement and convincing me to apply, and all the PhD programmes that rejected me, making me wait for this perfect opportunity.

In addition to my colleagues and work friends, I want to thank my friends and family for their support during these five years. My Parisian friends, I am so happy that complaining about the French government has bound us forever. Thank you for all the fun weekends and the video calls during the pandemic. Manali, Mona, thanks for always being there to babble about politics, academia, health care, and psychology. My friends from the Vosges, les gens, who can believe it is already 11 years since we left high school? I am so grateful I still get to hang out with you, seeing you is like taking a deep breath and reminding me of what really matters. My Erasmus friends, Martina, Dilek, Jan, Tessa, no matter how long it has been since we last saw each other I know I have you with me. You know how hard it is to have your heart split in many countries.

Papa et maman, merci pour les bons gènes et environnement. Merci pour la persévérance, la rigueur et la détermination que vous m’avez transmises. Merci de toujours me soutenir dans mes aventures lointaines. Noé, quand j’ai commencé ma thèse tu étais seulement au lycée, maintenant c’est déjà à ton tour de réfléchir au doctorat. Merci de me faire sentir comme un bon exemple à suivre et me rappeler de l’être. Ophélie, merci de le faire rire et nous avec. Maurine, merci de nous rappeler que l’important c’est de faire ce qu’il nous plaît, où cela nous plaît. Papy Michel, mamie Francette, papy Daniel et mamie Cécile, les tontons tatans et les cousins, merci d’être toujours présent même si je ne reviens pas assez souvent vous voir. Céline, merci pour cette magnifique couverture et mise-en-page et pour t’être adaptée

à mon style de travail “à la dernière minute”. Liebe Familie Buschmeier, danke für die schöne Zeit, ich bekomme immer viel Energie, wenn ich euch treffe. Konstantin, thank you for always bearing with me. Thank you for choosing to stay in Amsterdam with me. As much as I do not wish for another lockdown, I could not have spent these years with anyone better than you. These five years went so fast, and I don’t think we realize how much we changed. Thank you for your patience and your steadiness, holding me afloat in the changing weather. Thank you for making me laugh, making me rest, making me live.

Education and mental health shape individuals' wellbeing and their opportunities in life. Understanding the relationship between these two domains and their intergenerational transmission is valuable for researchers, policy makers, and society as a whole. With this doctoral dissertation, I tried to further our understanding of the causes of individual differences in educational outcomes, their relationship with mental health, and their co-intergenerational transmission. In my research I worked with observational data in large samples, triangulating across several genetically-informed designs to study genetic and environmental transmissions.