Nuffield Foundation

Nuffield Council on Bioethics

February 2025

Navigating genomics and education: insights, opportunities and challenges

Contents

Acknowledgements

This report was lead authored by Emma Meaburn, with substantive contributions from Natalie Michaux.

We would like to thank colleagues at the Nuffield Foundation and Nuffield Council on Bioethics who provided valuable support and oversight, including Josh Hillman, Danielle Hamm, Martin Davies, Hayleigh Simpkins and Sarah Walker-Robson. We are also grateful to members of the Nuffield Council on Bioethics for their feedback on the report and particular thanks go to the chair, Sarah Cunningham-Burley, for her reflections and insights throughout the drafting process.

We would also like to thank all the following external experts who contributed to this report by providing input, guidance, information, or critical review:

Professor Anna Vignoles

Professor Daniel Ansari, Department of Psychology & Faculty of Education, Western University

Dr Daphne O. Martschenko, Stanford Center for Biomedical Ethics

Professor Denis Mareschal, Centre for Brain and Cognitive Development, Birkbeck, University of London

Professor Emily Jones, Centre of Brain and Cognitive Development, Birkbeck, University of London and Institute of Psychology, Psychiatry and Neuroscience, King's College London

Professor Eric Turkheimer, PhD. Hugh Scott Hamilton Professor, University of Virginia

Professor Jean-Baptiste Pingault, Department of Clinical, Educational and Health Psychology, University College London

Professor Kathryn Paige Harden, Department of Psychology, University of Texas at Austin

Professor Maggie Snowling, University of Oxford

Professor Michael Thomas, Centre for Educational Neuroscience, Birkbeck, University of London

Professor Michelle Luciano, Psychology, University of Edinburgh

Professor Robert Plomin, CBE FBA FMedSci, Social, Genetic and Developmental Psychiatry Centre, King's College London

Professor Sarah-Jayne Blakemore, FBA FMedSci FRS, University of Cambridge

Professor Sophie von Stumm, Department of Education, University of York

Professor Terrie E. Moffitt, MBE, Department of Psychology & Neuroscience, Duke University and Institute of Psychiatry, Psychology & Neuroscience, King's College London

Dr Tim Morris, Centre for Longitudinal Studies, Social Research Institute, University College London

Professor Yulia Kovas, King's College London (visiting professor)

About the Nuffield Council on Bioethics and the Nuffield Foundation

About the Nuffield Council on Bioethics (NCOB)

The NCOB is a leading independent policy and research centre that identifies, analyses and advises on ethical issues in biomedicine and health to benefit people and society.

Established by the Trustees of the Nuffield Foundation in 1991, since 1994 it has been funded jointly by the Nuffield Foundation, the Medical Research Council and Wellcome.

For over 30 years, we have identified and tackled some of the most complex and controversial issues facing societies across the globe.

Find out more:

Website: www.nuffieldbioethics.org

LinkedIn: Nuffield Council on Bioethics

About the Nuffield Foundation

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice.

The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory.

Find out more:

Website: www.nuffieldfoundation.org

X and BlueSky: @NuffieldFound

LinkedIn: Nuffield Foundation

How to cite this report

Meaburn, EL. (2025) Navigating genomics and education: insights, opportunities and challenges, Nuffield Council on Bioethics and Nuffield Foundation

Executive summary

This scoping report is the first publication from a joint project between the Nuffield Foundation and the Nuffield Council on Bioethics. It outlines key findings and emerging directions in educational **genomic** research, and what is understood about the processes linking genetic differences to variation in **phenotypes** (measurable characteristics) related to education.

We examine how these findings are shaping the direction of further research and how close they are to being applied in real-world contexts, highlighting the scientific and practical challenges, and touching on some of the ethical issues that arise. This report therefore provides a foundation for further exploration of the ethical and policy implications of applying genomic insights to education, as well as the research gaps identified.

Approach

To inform this report, we carried out a scoping exercise using desk-based research, reviewing academic and grey literature, and holding semi-structured discussions with key researchers who were identified based on expertise and publications. Discussions used open questions to gather perspectives on key scientific, ethical and practical insights.

Key scientific insights

Individual differences in social and behavioural phenotypes, including those relevant to education such as achievement and mental health, are a consequence of both **genetic variation** and environmental factors (which include social and familial factors), and their interplay across development.

Advances in technologies that enable rapid, low-cost **DNA sequencing** and genotyping, alongside the growing availability of large-scale genomic datasets, have made **genome-wide association studies** (GWASs) possible. These studies have identified thousands of genetic variants associated with educationally relevant phenotypes. Among these, years spent in education is the most studied social phenotype in genomics due to its relative ease of measurement and its strong

associations with health and economic outcomes. Other educationally relevant phenotypes have been examined using GWASs, but many remain under-examined.

Although specific genetic associations have been identified, social and behavioural phenotypes linked to education are collectively influenced by thousands of genetic variants and are therefore highly **polygenic**. As a result of polygenicity, while GWASs have identified numerous genetic associations, much of the genetic contribution to educationally relevant phenotypes remains unexplored. GWAS research has also revealed that genetic variants associated with one phenotype often link to multiple other phenotypes – a phenomenon known as **pleiotropy**. This helps in understanding why some learning, behavioural and mental health difficulties overlap or co-occur. However, polygenicity and pleiotropy can make it difficult to interpret genetic associations and understand the mechanisms linking genetic variation to differences in educational phenotypes.

Environments are themselves also influenced by genetic variation, as an individual's DNA sequence can have an impact on the environments they experience, have access to, or select. This overlap between genetic and environmental influences – known as **gene–environment correlation** – complicates efforts to disentangle their respective contributions to educational disparities and to determine their relative impacts.

Data from GWASs have also led to the development of **polygenic indices** (PGIs). Polygenic indices combine the effects of thousands of genetic variants, each with a very small impact, to account for some population-level variation without requiring additional data. A PGI for number of years in education has been developed, and accounts for a proportion of variation comparable to some demographic measures (such as household income), though it predicts less variation than prior academic achievement. PGIs for other educationally relevant phenotypes have also been created, but these typically account for much smaller proportions of individual differences.

Social and practical challenges

The complexity of genetic influence on educationally relevant phenotypes means that interdisciplinary collaboration across the sciences, humanities and social sciences is needed to design, conduct and interpret research findings. Collaborative approaches will also assist with accurate and responsible communication of research findings. Responsible scientific conduct and communication is particularly important in educational genomics where misinterpretation and misuse of information can embed inequities, promote eugenic ideologies and have other wide-reaching sociopolitical ramifications.

Limited diversity in genomic datasets, and how population level descriptors are used in genomic research, also contribute to inequities, as well as misinterpretation of findings. The vast majority of GWAS participants are of European genetic ancestral descent (**genetic ancestry**), despite their making up a minority of the global population. Genomic findings are not readily extrapolated across diverse communities and

populations; if such findings are misinterpreted as being generally applicable, there are risks of drawing inaccurate – even harmful – conclusions about the sources of difference, and further reinforcing inequities experienced by marginalised and under-represented groups. New methods have been developed to allow researchers to utilise data from multiple populations simultaneously, and initiatives introduced to encourage diversification of genomic data. However, progress remains slow and such diversification attempts raise their own set of ethical questions, including around engagement, barriers to participation and data ownership.

Limited diversity is an issue in research design as well as in the data used. Beyond years spent in education, relevant phenotypes studied to date largely consist of mental health and some cognitive measures in adult populations. Greater diversity in, and depth of, phenotypes examined beyond these – alongside a fuller range of social and environmental factors – are therefore needed for deeper and more holistic insights into differences in educational experiences and trajectories.

Other contextual issues arise when considering the potential for translation of genomic insights into education practice and policy. Some of these are posed by gaps in knowledge and methodological limitations. Some limitations arising from the complexity of polygenicity and pleiotropy, for example, are not surmountable, as they are characteristic of the nature of genetic influence. Others can be addressed, at least in principle. Greater diversity in genomic and environmental data to address some of the extant biases, along with greater understanding of the underpinning social and neurobiological mechanisms behind individual differences in learning and brain function, may be achieved by further interdisciplinary research. In turn, this greater understanding may help to develop a wider range of PGIs with increased predictive power, which could form the basis of more robust and equitable PGI-based predictive tools. The timeframe for overcoming challenges posed by knowledge gaps and limitations is, however, uncertain, and each challenge raises its own set of ethical considerations which require in-depth deliberation.

PGIs are a specific example of where questions around translation currently exist. In the research context, PGIs can help researchers to untangle the roots of educational disparities across generations; investigate whether children most likely to profit from a particular environment can be identified; and explore causal processes linking certain environments to educational outcomes. However, outside of research contexts, they have received less attention, and the scientific and ethical questions raised by real-world translation represent a significant gap in extant knowledge. From an ethical perspective, the challenges arising from translation into practice remain underexplored, including the fundamental ethical question of whether PGIs should be translated into education at all. From a practical perspective, one particular issue needing further exploration is the imprecise nature of PGIs for making individual-level predictions about educational outcomes, and the potential impacts arising from attempts to do so. This makes their value and translation into policy and practice unknown. A greater understanding of the safeguards or boundaries that may render translation practically and ethically acceptable would likely be of benefit not only to

educators and policymakers, but also to the scientific community in developing the scope of future genomic and genetically-informed research.

Growing awareness of, and access to, genetic testing by the general public also serves to complicate the landscape. An increase in demand could serve to accelerate the translation of PGIs before the scientific and ethical questions have been considered, and before appropriate policy has been developed to guide educators. Although genomic literacy in the general population is low, PGIs are increasingly available commercially via direct-to-consumer genetic testing companies, including for social and behavioural phenotypes relevant to education. Further exploration on the potential ethical and practical impacts of greater supply and demand of PGIs on education practice and policy will therefore be of benefit.

It is, however, unknown how much awareness of and support for PGI-based educational interventions currently exists in educational practice – still less how much capacity there is to facilitate such interventions in already resource-limited education systems. Further exploration of whether education has the understanding, appetite or resources to utilise genomic information, and the implications of these for policy, is therefore essential.

1 Introduction

Genomic research has progressed rapidly in a short space of time. The year 2004 marked the publication of the near-complete **human genome** sequence,[1] and in the intervening two decades scientists have moved beyond understanding what the average human genome looks like to cataloguing patterns of genomic variation – differences in DNA sequence between individuals – across human populations.[2]

These advances, combined with technological and computational developments, have enabled researchers to identify specific genetic variants associated with **phenotypes** that impact or relate to education, providing novel insights into their genomic architecture. They have also led to the development of **polygenic indices** (PGIs), which index some of the genetic influences on a phenotype for an individual by summing up the cumulative effect of multiple genetic variants which individually have a very small impact. PGIs are now available for some educationally relevant phenotypes, including how long one spends in education, cognitive function and behavioural difficulties. However, it is important to note that educationally relevant phenotypes are shaped by both genetic and environmental variation and their interplay across development, making it challenging to establish clear causal links.

This scoping paper is the first publication in a broader collaborative programme of work between the Nuffield Foundation's Education domain and the Nuffield Council on Bioethics (NCOB), examining research in **genomics** and neuroscience and its social and ethical implications for education policy and practice. It provides an overview of key concepts, genomic methods and research insights, and the wide range of ethical issues raised by these developments and potential applications. How long a person spends in education is predictive of important social, economic and health outcomes,[3] so understanding genomic – and environmental – contributions to educational disparities is crucial for addressing inequity in society. This report is

1 International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome *Nature* **431**: 931–45.

2 The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation *Nature* **526**: 68–74.

3 Farquharson C, McNally S, and Tahir I (2024) Education inequalities *Oxford Open Economics* **3(1)** i760–820.

primarily scoping the scientific terrain as a first step to developing a fuller consideration of the ethical context of both the research and its possible applications. Genetic and genomic research raises a wide range of ethical issues, and the area of genomic prediction especially, as highlighted in the recent NCOB reports on **health-related prediction** (a collaboration with the Ada Lovelace Institute) and on **working towards a gold standard of ethics in genomic healthcare and research**. This contentious area requires both an understanding of contemporary scientific practices and possibilities, alongside considered ethical deliberations.

The report pays particular attention to the use of PGIs to untangle the genetic and environmental contributions to social and behavioural differences. It also considers the scientific utility of using PGIs as tools to identify children who might benefit from early access to tailored educational support. Identifying educational and learning vulnerabilities before they arise – as a way to help support individuals – has clear and understandable appeal, but it raises profound scientific and ethical questions. While PGIs can sometimes explain a modest proportion of population-level variation in a phenotype, they – like all predictors – are probabilistic and not accurate for predicting individual-level outcomes.

Beyond the scientific and ethical debates about the potential of PGI-based individual-level prediction for educational outcomes, PGI reports are increasingly available and accessible in the direct-to-consumer (DTC) market. This raises concerns about how such tools might be harnessed and applied in real-world settings, and the potential they hold for reinforcing – or creating new forms of – inequities in accountability-driven educational systems.

We hope this report will be a useful resource for current and future researchers in genomics and education, as well as for policymakers and educational practitioners, and others who want to engage in considering the implications of these developments.

1.1 The scope of the report

This report examines the drivers of, and key insights from, **molecular genomic research** into phenotypes that relate to or impact education. These include time spent in education, achievement, general cognitive ability ('intelligence'), specific learning difficulties, mental health conditions and personality traits. For simplicity, we use the term '**phenotype**' to refer to these factors and traits throughout the report, while recognising knowledge of the phenotype is limited to the measures available to researchers.

While the report touches on **twin study** research, its primary focus is on **molecular genetic research** that identifies inherited DNA differences associated with variation in phenotypes, and how this knowledge has led to the creation of **PGIs**[4] for educationally relevant phenotypes. PGIs represent something of a paradigm shift, moving the focus from estimating net genetic influences in populations to estimating an individual's genomic likelihood, compared with others in the population, of exhibiting a specific phenotype.

Like other predictors of educational phenotypes, PGIs account for some of the variation observed between individuals but are imprecise tools for making predictions at the individual level. Also, the mechanisms and processes through which PGIs influence specific phenotypes remain largely unknown, and their predictive accuracy diminishes when applied to populations different from those studied in the original 'discovery' research. As a result, many scientists consider PGIs to be useful primarily at the population level, where they can serve as research tools to tackle genetic **confounding** and to explore the nature of relationships between environmental and social factors and differences in educational outcomes.

Outside of these academic discussions, however, PGIs are becoming increasingly available and accessible from DTC companies.[5] It is plausible that a PGI for time spent in education – or another social or behavioural phenotype relevant to education – could soon be offered directly to consumers and used in ways that are inappropriate or even harmful.[6] This prospect is a further motivator for our work to map the scientific and ethical challenges across genomics and neuroscience in education so that societal and individual harms and benefits can be thoroughly understood.

It is important to acknowledge upfront the limitations of phenotypes used in educational genomic research. Phenotypes such as educational attainment, achievement and

4 In line with recommendations, we use the term 'PGI' instead of 'polygenic score' or 'polygenic risk score' so as not to imply a value judgement where one is not intended. See Becker J, Burik CAP, Goldman G, *et al* (2021) Resource profile and user guide of the Polygenic Index Repository *Nature Human Behaviour* **5**: 1744–58.

5 Park JK and Lu CY (2023) Polygenic scores in the direct-to-consumer setting: challenges and opportunities for a new era in consumer genetic testing *Journal Personalized Medicine* **13(4)**: 573.

6 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

general cognitive ability represent only one dimension of schooling, are measured with error and in some instances may be a proxy for the 'real' phenotype of interest. For example, in genomic research, 'number of years of formal schooling' is often used as a proxy for educational attainment, a metric that does not necessarily reflect achievement or ability; that overlaps with socioeconomic factors which may not have been controlled for; and in some respects can be considered as an input rather than an outcome measure.

This report does not explicitly seek to delineate ethical concerns that apply specifically to educational genomic research or genomic research more generally, such as the collection, storage and management of genomic data, privacy and consent.[7] Some of these concerns have been outlined in other NCOB reports and remain a key part of our ongoing work on genomics and neuroscience in education. While it is impossible to ignore these ethical concerns, this report provides an initial overview of some key issues, with more in-depth consideration to follow in future work. The report aims to provide a valuable foundation for informed, evidence-based conversations among stakeholders about the proposed applications of PGIs in research and in practice, and the ethical issues therein. Our hope is that by laying out the scientific complexities and limitations in this way, and by highlighting some of the key ethical considerations, the report will be useful to researchers, educational professionals and policymakers, and will guide discussions on whether genomic information for educationally relevant phenotypes might be used responsibly in the real world.

1.2 How the report was informed

To inform this report, we carried out a scoping exercise using desk-based research, reviewing academic and grey literature, and holding semi-structured discussions with key academic researchers who were identified based on research outputs. The researchers identified had a diverse range of perspectives on how current knowledge of genomic contributions to social and behavioural phenotypes may impact educational research, practice and policy. The researchers and their affiliations are listed at the end of the report.

The discussions used open questions to gather perspectives on the key scientific and ethical insights.

7 Oliva A, Kaphle A, Reguant R, *et al* (2024) Future-proofing genomic data and consent management: a comprehensive review of technology innovations *GigaScience* **13**: giae021.

1.3 Report structure

The report is written to be read sequentially, but readers may choose to focus on specific sections that are more relevant to them. Each section contains key points at the start. Ethical questions raised by the use, interpretation and application of genomic data are highlighted throughout.

As far as possible, we have minimised the use of scientific terminology and technical terms. However, some terms are necessary as they have a very specific meaning. When these terms appear first in each section, they are indicated in **bold** and explained in the accompanying glossary.

For the interested reader, we also provide more detailed information on DNA (**Annex 1**) and DNA sequence variation and genotyping (**Annex 2**), **heritability** (**Annex 3**), **genome-wide association studies (GWASs)** (**Annex 4**), and polygenic indices (**Annex 5**). These annexes are designed to be stand-alone and can be skipped if preferred.

Section 2 introduces the basics of **DNA**, **genetic variation**, GWASs and PGIs. It also explains the scientific rationale behind GWASs and how associations between genetic variants and phenotypes should be interpreted.

Section 3 summarises the main drivers of discovery research, the current knowledge of specific genomic contributions to educationally relevant phenotypes, and ethical issues identified as a result of these advancements.

Section 4 explores the opportunities and challenges in the translation of educational genomic research findings. It outlines how PGIs are being used as a methodological tool to investigate the interconnected influences of genes and environment, and to examine causal processes linking genomic and environmental factors to educational outcomes. It also examines the rationale for using PGIs as predictive tools to identify individuals at higher or lower likelihood of social or behavioural difficulties.

Section 5 considers the implications of PGIs for research, education practice and policy, highlighting key scientific, practical and ethical challenges across contexts, identifying areas where further exploration is needed both to fill gaps in knowledge and to properly consider whether (and in what circumstances) translation into policy and practice might be ethically undertaken.

2 DNA, genetic variation and genomic research methods

Key messages

- For **complex phenotypes**, individual differences are a result of a multitude of both genetic and environmental influences.
- **Genome-wide association studies** (GWASs) can identify genetic variants that correlate with individual differences in a **phenotype** of interest.
- The results of a GWAS are sensitive to the characteristics of the population studied, including patterns of **genetic variation** and environmental factors, as well as how the phenotype is measured.
- The results of a GWAS can be used to create individual scores, called **polygenic indices** (PGIs), which estimate some of an individual's genomic liability towards a given phenotype.
- Genetic associations, and by extension PGIs, are not immune to **confounding** and can pick up influences from other factors that contribute to human variation, including the environment.
- The nature of the associations between genetic variation and a given phenotype are generally very difficult to determine.
- Genetically homogeneous populations from western societies of limited diversity are over-represented in GWASs. As a result, insights from current GWASs do not apply to all people equally.

One of the primary goals of **genomic** research is to explore whether and how genetic variation contributes to differences in a phenotype within a population. In this section we outline some of the key molecular methods and statistical tools used in genomic research, along with their challenges, and explain how the results of GWAS analyses are used to construct PGIs. We expand on some common misconceptions in the interpretation of genomic and genetically informed research,

Continued >>

and on why the lack of diversity in genomic research poses a scientific and ethical barrier to translation.

This section serves as a technical primer. Readers already familiar with DNA, genetic variation, GWASs and PGIs may wish to skip ahead to **Section 3**. For those less familiar with these topics – particularly for those seeking to better interpret and engage with findings from genomic research as it relates to education – we hope that this section provides a useful introduction.

2.1 A note on ethical vigilance

Genomic research into individual differences in social measures and human behaviour, such as intelligence and personality, is controversial.[8] This is rooted partly in the historical development and use of genetic research in the early twentieth century, when eugenic social policies and laws were used to justify human rights violations and even genocide, all aimed at 'improving' the human gene pool and eradicating so-called undesirable traits and people.[9]

We acknowledge the legacy of these practices and their continued importance in shaping contemporary educational genomic research. This underscores the need for critical reflection and ethical vigilance to avoid repeating past mistakes. In this context, educational genomics requires ethical scrutiny to ensure that research promotes fairness and equity, and that findings are not misinterpreted or used – deliberately or not – to cause harm to individuals, groups or society.[10] As outlined later in this report, risks include genetic discrimination, stigmatisation and fuelling harmful practices or ideologies.[11]

While the report focuses on the genetic contributions to individual differences in social and behavioural phenotypes related to education, genetic contributions are not the only source of these differences. A particular concern identified during the writing of this report is that focusing too much on genomic factors may risk overlooking the importance of social, cultural and environmental factors in influencing disparities in educational outcomes, and the recasting of structural and historical inequities as biological characteristics of individuals.

8 Comfort N (2018) Genetic determinism rides again *Nature* **561(7724)**: 461–63.

9 Kevles DJ (1995) *In the name of eugenics: genetics and the uses of human heredity* (Cambridge, MA: Harvard University Press).

10 Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–49.

11 Ibid: Meyer *et al* consider the broader risks of socio-behavioural genomic research; see also the 2023 NIH NHGRI roundtable on the topic.

However, DNA differences are part of the story, so they are relevant to anyone concerned with, or impacted by, educational disparities. If we ignore the role of genes, we risk overlooking a source of differences between students that could lead to supportive educational interventions and practices.

2.2 DNA, genes and genomes

We are built of cells – trillions of them. Located in the nucleus of each cell is a set of genetic instructions called a **genome** that determines the function and activity of the cell. The genome is comprised of a chemical called deoxyribonucleic acid, or **DNA**, which consists of 6.2 billion 'letters'. These letters (or 'bases') are adenine (A), cytosine (C), guanine (G) and thymine (T), and they make up the steps on the spiral staircase of the double helix of DNA. The order of the As, Cs, Gs and Ts in the genome – or **DNA sequence** – is very important, as it guides the biological processes and reactions that are essential for the structure and function of cells (see **Annex 1**).

Many of these processes happen as a result of the information contained in the DNA sequence being translated, via gene expression, into proteins such as hormones, neurotransmitters and enzymes. The **human genome** contains around 20,000 of these protein-encoding stretches of DNA sequence, or **genes**. However, most of the human genome (around 98.5%) does not encode for proteins. This is called **non-coding DNA**. The function of non-coding DNA is not yet fully understood, but it is known to play a role in regulating how and when the information contained in DNA is accessed and used by cells.

2.3 Genetic and genomic variation

Importantly, there are differences in the DNA sequence between people in a population; if you were to compare the DNA sequence of any two apparently unrelated individuals, you would find that on average their genome will **differ at around 27 million positions**, or 'letters' (0.4%). The DNA differences between individuals are referred to as 'genetic variation' and occur throughout the genome. In this report we use the term genetic variation to refer to differences in DNA sequence in a specific region (or 'locus') in the genome, and 'genomic variation' to refer to differences spread throughout the genome.

When two or more versions of DNA sequence exist at a given location in the genome in a population, this is called an **allele**. As humans inherit two copies of DNA (one from each parent), the combination of the maternally and paternally inherited alleles for a genetic variant can be the same (homozygous) or different (heterozygous). The combination of alleles at a given locus is called a **genotype**.

One of the most widely studied types of genetic variation in genomic research is **single nucleotide polymorphisms** (SNPs). These are single-base (A, C, G or T) changes in the DNA sequence (see **Annex 2**).

2.4 Estimating the heritability of complex phenotypes

How important is genomic variation in explaining the differences observed among people within a population? Specifically, to what extent does the cumulative impact of genomic variation – across all the cells in a person's body – account for differences in behaviour and functioning? How does this genomic influence compare with environmental factors, such as the neighbourhood, home environment or broader social conditions?

Rare and severe learning and developmental conditions with clear-cut inheritance patterns within families (termed **Mendelian** or **monogenic conditions**) are caused by a deleterious change in DNA sequence. Such a change typically disrupts the function of a gene and, under usual environmental conditions, is both necessary and sufficient for the condition to occur. For example, **Phenylketonuria (PKU)** is a rare metabolic disorder resulting from a mutation in the PAH gene. If untreated, PKU leads to a neurotoxic build-up of phenylalanine, causing impaired cognitive development in affected individuals.

However, even for rare and severe monogenic conditions such as PKU, outcomes can be significantly improved through environmental interventions. In the UK, PKU is routinely tested for, and newborns who test positive can be placed on a life-long modified diet low in phenylalanine, which supports typical development.

The one-to-one relationship between genotype and outcome can make it easier to identify the genetic origins of monogenic conditions, especially with the use of **DNA sequencing** technologies and family data (data on both biological parents and their child).[12] Consequently, many rare genetic variants responsible for monogenic forms of severe intellectual disability, developmental delay, and severe speech and language conditions have been identified.[13] In many cases, the condition is caused by a *de novo* mutation that occurs spontaneously in the affected child.

In this report, we focus on complex social and behavioural phenotypes related to education that arise from both genomic and environmental influences and their interplay. What is currently known about the relative contribution of genomic variation to differences in educationally relevant phenotypes among individuals? Specialised **twin studies** and **adoption studies** address this question by comparing the resemblance between individuals with varying degrees of genetic and environmental relatedness to estimate the relative contributions of nature (genomic variation) and nurture

12 Ionita-Laza I, Makarov V, Yoon S, *et al* (2011) Finding disease variants in Mendelian disorders by using sequence data: methods and applications *American Journal of Human Genetics* **89(6)**: 701–12.

13 Deciphering Developmental Disorders Study (2015) Large-scale discovery of novel genetic causes of developmental disorders *Nature* **519(7542)**: 223–8.

(environmental variation). While twin studies make certain assumptions,[14] they consistently show that nearly every aspect of human individual differences are heritable – that a portion of the differences measured between people can be attributed to DNA differences.[15] The remaining variation is attributed to environmental factors (everything other than DNA) that are shared and non-shared by twins.[16]

It is important to be aware that **heritability** estimates the relative contribution of genetic influences to a phenotype in a specific sampled population at a specific sampled time. As a result, it does not have a fixed 'true' value and will vary across contexts, such as different cultures, educational systems, communities and time frames. See **Annex 3** for more information about heritability.

Additionally, demonstrating that a phenotype is heritable does not imply that a person's outcome is determined by their genetic make-up ('genetic determinism'): contextual environmental factors are always important. Nor does heritability mean that an outcome is inevitable and unavoidable ('genetic fatalism'), as changes in the environment may alter the expression of a heritable phenotype. For example, the rise in short-sightedness from 20% to 80% in a single generation illustrates how population-wide environmental shifts can influence the expression of a highly heritable phenotype.[17] Targeted environmental interventions – such as wearing glasses – can mitigate the impact of this heritable condition. **Annex 3** provides further guidance on the interpretation of heritability estimates.

Methodological advances have recently made it possible to estimate the heritability of a phenotype using measured genomic variation in unrelated samples, bypassing the need for specialised twin and adoption cohorts. This is called SNP-based heritability. It provides an estimate of the proportion of phenotypic variation that can be attributed to measured genomic variation in a sample (GWAS; see **Annex 3**).

While twin and adoption studies investigate the net effects of genes to provide estimates of heritability, they do not reveal the specific genetic variants involved, how many contribute or how they function.[18] To explore these questions, **molecular genomic research** is required that measures genomic variation directly, and links it to differences in biological processes, the brain and phenotypes of interest.

14 The 'equal environments' assumption is that identical and non-identical twins experience similar environmental exposures, meaning that identical twins behave more alike than non-identical twins due to their greater genetic similarity.

15 Polderman TJ, Benyamin B, de Leeuw CA, *et al* (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics* **47(7)**: 702–9.

16 Plomin R (2011) Commentary: why are children in the same family so different? Non-shared environment three decades later *International Journal of Epidemiology* **40(3)**: 582–92.

17 Tedja MS, Haarman AEG, Meester-Smoor MA, *et al* (2020) The genetics of myopia, in *Updates on myopia*, Ang M and Wong T (Editors) (Singapore: Springer), pp95–132.

18 Friedman NP, Banich MT, and Keller MC (2021) Twin studies to GWAS: there and back again *Trends in Cognitive Science* **25(10)**: 855-69.

2.5 Genome-wide association studies

Following the completion of the human genome sequencing project,[19] international efforts focused on cataloguing patterns of genomic variation across human populations.[20] The characterisation of genomic variation allowed researchers, for the first time, to test whether differences in genotype between people are associated with differences in complex phenotypes.[21]

As a result, the past 15 years have seen a surge of molecular genomic research aimed at identifying specific genetic variants that contribute towards educationally relevant social and behavioural phenotypes.

Early efforts focused on examining genetic variation in a handful of **candidate genes** that were hypothesised to play a biological role in the phenotype of interest – such as neurotransmitter genes in relation to intelligence. However, candidate gene studies of complex phenotypes have failed to replicate consistently, and reported associations are now assumed to be false positives.[22] This failure is attributed primarily to publication biases and the incorrect assumption that genetic variants of moderate to large effect underlie complex phenotypes.

More recently, technological and computational advances have made it possible to move away from a single-gene approach. Researchers can now systematically scan across the human genome to test if common genetic variants – typically SNPs – associate with a specific phenotype. These studies are called genome-wide association studies (GWASs)[23].

In a GWAS, researchers collect DNA samples from very large groups of unrelated (population-based) or related (**family-based**) individuals. The DNA of each person is collected and the DNA sequence read directly or inferred using statistical methods at often millions of positions across the genome. This is done using **SNP microarrays** or **DNA sequencing** technology such as **whole-exome sequencing** or **whole-genome sequencing**. More detail on GWASs can be found in **Annex 4**.

The genotype data generated in this way undergo careful quality control steps to aid the reliability of results, and each genetic variant is assessed for statistical association with the phenotype. Variants that reach a predetermined analytical threshold indicate a region of the genome (which may include several genetic variants) that

19 International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome *Nature* **431**: 931–45.

20 The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation *Nature* **526**: 68–74.

21 Kruglyak, L (2018) The road to genome-wide association studies *Nature Reviews Genetics* **9**: 314–18.

22 Chabris CF, Hebert BM, Benjamin DJ, *et al* (2012) Most reported genetic associations with general intelligence are probably false positives *Psychological Science* **23(11)**: 1314–23.

23 Uffelmann E, Huang QQ, Munung NS *et al* (2021) Genome-wide association studies *Nature: Reviews Methods Primers* **1**: 59.

correlates with the phenotype. This is called a genetic association, and the associations identified are replicated in independent samples to ensure that they are reliable.

The ability of a GWAS to detect genetic associations depends on several factors. Firstly, it depends on how frequent the genetic variant is in the population being studied. If the change in DNA sequence arose many generations ago and is not deleterious to health, the variant can become common in the population. If the change arose more recently or is detrimental to survival and fecundity, then it is likely to be rare, perhaps even specific to a family or a person. Rare variants can have a large effect on an individual, but as they are infrequent, they have a small impact on variance in the population. This makes them harder to detect and quantify in a GWAS. Secondly, it depends on how big an impact the genetic variant has on a phenotype. For example, a genetic variant that accounts for a difference of two years of schooling in a population would have a large **effect size**, while one that accounts for two weeks would be considered very small. In fact, most common genetic variants identified in a GWAS have extremely small effect sizes, meaning each variant predicts only a tiny fraction of the variation in the phenotype.

The phenomenon of many common genetic variants of small effect size collectively influencing a phenotype is known as **polygenicity**. Polygenic variants become easier to detect in GWASs with very large sample sizes, i.e. upwards of 100,000 individuals.

The results of a GWAS are stored as **GWAS summary data**.[24] These data show the strength of the evidence for an association with a phenotype at a group level (measured by a **p-value**), the effect size, and the direction of the link (i.e. raising or lowering the value of the phenotype) for every genetic variant tested.

Freely accessible data resources have been developed that collate and catalogue published GWASs, including the central **EMBL-EBI GWAS Catalog**, the **IEU Open GWAS project** and the **Global Biobank Engine**. These resources are fully searchable and make summary data for tens of thousands of GWASs available to researchers and other interested parties.

2.6 Genetic ancestry and lack of genomic diversity

To minimise the risk of spurious genetic associations arising from **population stratification**, GWASs have generally been performed using participants who share similar patterns of genomic variation.[25] Researchers typically group these genetically similar individuals using broad continental descriptors such as 'European', 'Asian' or 'African' and refer to these groupings as **genetic ancestry**. However, this approach fails to capture the complexity of human genetic diversity, as ancestry is a continuum

24 Ibid.

25 Ibid.

that varies even within a group.[26] While new methods now allow the inclusion of individuals from diverse populations, the vast majority of GWASs published so far have been performed in populations of European genetic ancestry; recent estimates indicate that 94.6% of all discovery GWAS participants are of European genetic ancestral descent, despite this group representing only 16% of the global population.[27]

This bias presents a serious challenge because, due to differences in environmental contexts, the frequency of genetic variants, their effect size and patterns of correlation with each other, GWAS findings are not directly comparable or perfectly 'portable' across populations. As a result, genomic insights cannot be meaningfully extrapolated from one population to another: to do so risks inaccurate or even harmful conclusions about the sources of differences among people. Due to the Euro-centric bias, these risks would disproportionately affect already marginalised and underserved groups and individuals.[28]

Moreover, while race and genetic ancestry descriptors can sometimes overlap, the two are not interchangeable – race is a **social construct**. There is widespread concern within and beyond the scientific community that without careful communication, genomic research findings **may be misinterpreted as evidence for genetically discrete racial groups**. This could perpetuate the harmful and incorrect notion that group-level disparities in health, education and behaviour are driven by genetic differences, further reinforcing damaging stereotypes.[29] In light of these concerns, the National Academies of Sciences, Engineering and Medicine (NASEM) recently released a report making recommendations on the use of population descriptors in genetics and genomics research,[30] to which journal editors are paying close attention.[31]

Where does this leave us? The lack of diversity poses a major barrier to the translational potential of genomics across all fields of inquiry, including education. However, 'roadmaps' are being developed to address this,[32] and the research community and funding bodies are driving numerous initiatives to tackle the

26 Lewis ACF, Molina SJ, Appelbaum PS, *et al* (2022) Getting genetic ancestry right for science and society *Science* **376**: 250–2.

27 Mills MC and Rahal CA (2019) A scientometric review of genome-wide association studies *Communications in Biology* **2**: 9.

28 Martin AR, Kanai M, Kamatani Y, *et al* (2019) Clinical use of current polygenic risk scores may exacerbate health disparities *Nature Genetics* **51**: 584–91.

29 Cerdeña JP, Grubbs V, and Non AL (2022) Genomic supremacy: the harm of conflating genetic ancestry and race *Human Genomics* **16(18)**.

30 National Academies of Sciences, Engineering, and Medicine (2023) *Using population descriptors in genetics and genomics research: a new framework for an evolving field* (Washington, DC: National Academies Press).

31 Feero WG, Steiner RD, Slavotinek A, *et al* (2024) Guidance on use of race, ethnicity, and geographic origin as proxies for genetic ancestry groups in biomedical publications *JAMA* **331(15)**: 1276–8.

32 Fatumo S, Chikowore T, Choudhury A, *et al* (2022) A roadmap to increase diversity in genomic studies *Nature Medicine* **28(2)**: 243–50.

challenge. For instance, the UK's **Our Future Health biobank**, the USA's **All Of Us Program**, and **H3Africa** have been established to work to ensure that global populations are better represented in genomic research, though their focus is often on health and not education. Analytical methods such as 'trans-ancestry' or 'cross-ancestry' GWASs[33] now allow researchers to consider multiple genetic ancestries simultaneously. These approaches are being facilitated by efforts to coordinate global biobank resources for large-scale GWASs.[34]

However, the diversification of genomic data remains slow and raises its own set of ethical questions.[35] Concerns include exploitative 'helicopter research' practices, where researchers from high-income countries exploit collaborations with low- and middle-income countries; a lack of meaningful community engagement; and a failure to prioritise the co-production of knowledge.[36] Structural issues have been identified, such as data governance, oversight of public–private partnerships, knowledge production and dissemination, and the role of funding bodies and journals in data diversification efforts.[37] There are also questions to be addressed about why minoritised groups may not want to participate in genomic research, and how the scientific community can work collaboratively with them to shape future research and governance mechanisms. One way in which some of these issues are being tackled is through the creation of stakeholder-led biobanks that collect and maintain ownership of genomic data, such as the **Native BioData Consortium**. It should also be noted that biases arise even within high-income settings, with many sections of society entirely unrepresented in the large-scale cohorts currently available.[38]

2.7 Beyond genetic associations

While GWASs seek to identify specific genetic associations, the results of a GWAS can tell researchers other useful things. For instance, SNP-based heritability is now routinely calculated as part of a GWAS analysis, as this can indicate whether there are genetic associations yet to be discovered and how much predictive power a **polygenic index** (PGI) might provide for the phenotype.

33 Peterson RE, Kuchenbaecker K, Walters RK, *et al* (2019) Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations *Cell* **179(3)**: 589–603.

34 Zhou W, Kanai M, Wu KH, *et al* (2022) Global biobank meta-analysis initiative: powering genetic discovery across human disease *Cell Genomics* **2(10)**: 100192.

35 Wellcome Trust (2024) Data and diversity (London: Wellcome Trust).

36 Martin AR, Stroud RE, Abebe T, *et al* (2022) Increasing diversity in genomics requires investment in equitable partnerships and capacity building *Nature Genetic* **54**: 740–5.

37 Hardcastle F, Lyle K, Horton R, *et al* (2024) The ethical challenges of diversifying genomic data: a qualitative evidence synthesis *Cambridge Prisms: Precision Medicine* **2**: e1.

38 Schoeler T, Speed D, Porcu E, *et al* (2023) Participation bias in the UK Biobank distorts genetic associations and downstream analyses *Nature Human Behaviour* **7**: 1216–27.

The data used to perform a GWAS, or the results of a GWAS analysis, can also be used to understand the genomic relationship (**genetic correlation**) between pairs of phenotypes.[39] An early GWAS observation was that genetic associations identified for one phenotype often correlate with other, sometimes apparently distinct, phenotypes. A potential explanation for these correlations is **pleiotropy**, where a genetic variant influences more than one phenotype.[40] Understanding the extent to which different phenotypes share common genetic influences is important, as it can help in developing testable hypotheses about shared pathways or processes, and clarifying why certain behavioural or learning conditions often overlap or co-occur.[41]

2.8 Distinguishing correlation from causation

The aim of a GWAS is to identify the causal effects of an individual's genotype on their phenotype. A genetic variant can be considered causal if, in a given **environment**, the individual's phenotype would have been different had the genetic variant inherited from their parents been different. However, the counterfactual (i.e. what would happen in an alternative scenario where only the DNA variant changed) cannot be tested, because researchers do not have experimental control over which genetic variants individuals inherit while keeping all other factors constant. This is a core challenge with naturalistic observational research, such as a GWAS, as it makes establishing causality difficult.

Since the DNA sequence a person is born with remains largely unchanged in most cells throughout their life course, confounding due to reverse causation can be ruled out for genetic associations. For instance, it is not possible for genetic variants associated with maths skill to be altered by performance in a maths test. This makes genomics distinct from other forms of analysis, where outcomes of interest (such as educational achievement) can potentially influence the environmental or behavioural risk factors (such as parenting practices or mental health status).

Because reverse confounding cannot occur with genetic associations, it has often been thought that classical confounding was not an issue, leaving only causal genetic effects. However, it is now becoming clear that environmental confounding can occur in observational genetic association analyses. This is because of the real-world phenomenon of **gene–environment correlation**, where an individual's DNA correlates with the environments they are exposed to or seek out. A variety of gene–environment correlation processes have been described that act within and between families

39 Hackinger S and Zeggini E (2017) Statistical methods to detect pleiotropy in human complex traits *Open Biology* **7(11)**: 170125.

40 Solovieff N, Cotsapas C, Lee P, *et al* (2013) Pleiotropy in complex traits: challenges and strategies *Nature Reviews Genetics* **14**: 483–95.

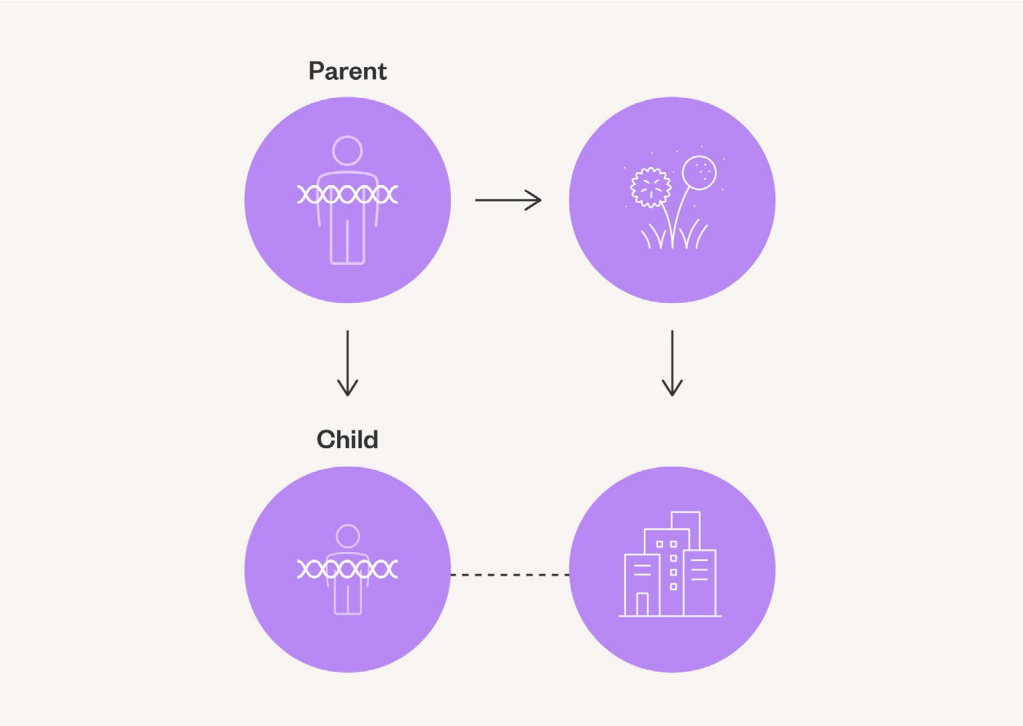
41 Watanabe K, Stringer S, Frei O, *et al* (2019) A global overview of pleiotropy and genetic architecture in complex traits *Nature Genetics* **51**: 1339–48.

which we do not cover in detail in this report.[42] The key point is that due to gene–environment correlation processes, genetic associations can also pick up the influences of social or demographic factors that contribute to the phenotype being examined.

For example, a genetic variant in a parent that causes allergies might prompt them to move from a rural to an urban area. The parent provides both their DNA and a rearing environment to their offspring. If the urban-born offspring inherits the allergy variant, a correlation is created between the variant and geography (i.e. the allergy-increasing allele becomes more common in urban areas than rural ones). A correlation is also created with any other offspring phenotype that is influenced by growing up in an urban setting.

As a result, if the offspring were to participate in a GWAS, a non-causal association may be detected between the allergy-related variant and pollution levels. While the variant has a causal effect on geography in the parent (i.e. where they chose to live) and is a valid predictor of pollution exposure in the offspring, it is not causal in the strictest sense that changing the variant in the offspring will alter environmental pollution levels. Instead, the association is induced because the parental genome is part-shared with the offspring and also shapes the rearing environment they provide. This is a form of **passive gene–environment correlation** (see Figure 1).

Figure 1: Schematic of a non-causal genetic association



42 Pingault JB, O'Reilly PF, Schoeler T, *et al* (2018) Using genetic data to strengthen causal inference in observational research *Nature Reviews Genetics* **19**: 566–80.

In the example shown in **Figure 1**, the genotype of a parent has a causal influence on allergies and their behaviour(s), such as where they choose to live. When this variant is inherited by the child, it will become non-causally correlated with the child's phenotype(s) that are a consequence of being raised in the environment that their parents created (shown by the dashed line).

Similarly, child educational outcomes can be influenced by parental genes through parental rearing behaviours and inherited social, economic and learning environments. These 'indirect' environmental effects can occur in addition to the 'direct' effects on child outcomes of inherited genetic variation. As discussed in **Section 4.4**, 'indirect' genetic associations appear to be more common for social outcomes and cognitive phenotypes that relate to education.[43] Other mechanisms, such as **assortative mating** (when individuals are more likely to choose a genetically similar individual to partner with, rather than randomly – a known phenomenon for education[44]) and population stratification can also produce non-causal genetic associations that complicate the interpretation of genetic associations.

As a consequence, estimates of genetic associations are likely to capture both the direct (putatively causal) effects of an individual's DNA, and the indirect effects of genetic relatives, assortative mating and population stratification. So, while genetic association is a first step in establishing a causal link between a genetic variant and a phenotype, it should be interpreted as correlational until additional evidence can be gathered supporting particular causal pathways.

2.9 Understanding neurobiology: from genetic variants, to genes, to biology

A primary motivation for conducting a GWAS for any complex phenotype is to gain novel insights into biology by implicating genetic variants in specific genes, biological pathways or processes.[45] In the context of education, a common aim is to improve researchers' understanding of the neurobiological mechanisms behind individual differences in learning and brain function.

In medicine, understanding biological mechanisms can inform drug discovery and development.[46] In education, neurobiological insights could support teachers in the

43 Howe LJ, Nivard MG, Morris TT, *et al* (2022) Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects *Nature Genetics* **54**: 581–92.

44 Mare, RD (1991) Five decades of educational assortative mating *American Sociological Review* **56(1)**: 15–32.

45 Uffelmann E, Huang QQ, Munung NS *et al* (2021) Genome-wide association studies *Nature: Reviews Methods Primers* **1**: 59.

46 King EA, Davis JW, and Degner JF (2019) Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval *PLoS Genetics* **15(12)**: e1008489.

classroom by informing the selection of effective pedagogical techniques and learning materials, or aid the development of effective intervention strategies for when – or even before – difficulties arise.

However, interpreting the biological implications of genetic variants identified through GWASs remains a significant challenge.[47] One reason is that – as mentioned – genetic associations can capture the confounding influences of the environment and other factors as well as the direct effects of an individual's biology. Even if this issue were resolved, the highly **polygenic** and **pleiotropic** nature of such phenotypes makes tracing pathways from genes to brain to phenotypes extremely complex. Additionally, associated variants tend to be inherited in 'chunks' (called '**linkage disequilibrium**'), making it difficult to identify the specific variant responsible for the association signal. These issues are further compounded by incomplete knowledge of biological pathways and networks in the brain, and how they develop across infancy to adulthood. For these reasons, translating GWAS findings into biological insights remains challenging – even in the medical sciences.[48]

Further issues can arise from the depth and mix of knowledge and skills required to interpret GWAS results. At a minimum, biological interpretation of GWAS data requires computational expertise and *in silico* analysis of additional biological datasets, such as proteomic, transcriptomic or epigenomic data.[49]

Ultimately, understanding the causal mechanisms underlying phenotypic variation requires basic scientific research that extends beyond GWASs. This includes studies using animal models and computational approaches, where shared neural, hormonal or developmental processes and mechanisms can be experimentally manipulated – such as through gene modification or altering environmental conditions – to observe their effects on learning and behaviour.

2.10 Polygenic indices for genomic prediction

Data from a GWAS can be used to calculate a genomic predictor of a phenotype known as a polygenic index (PGI).[50] A PGI is a person-specific value that summarises some of that individual's genetic predisposition to a phenotype. Instead of focusing on individual SNP-level genetic associations, PGIs sum up the effects of thousands of SNP-level genetic associations spread throughout the genome, weighted by their

47 Brandes N, Weissbrod O, and Linial M (2022) Open problems in human trait genetics *Genome Biology* **23**: 131.

48 Ibid.

49 Uffelmann E, Huang QQ, Munung NS *et al* (2021) Genome-wide association studies *Nature: Reviews Methods Primers* **1**: 59.

50 Becker J, Burik CAP, Goldman G, *et al* (2021) Resource profile and user guide of the Polygenic Index Repository *Nature Human Behaviour* **5**: 1744–58.

strength and direction of their GWAS association.[51] Details on how a PGI is calculated, with a simplified example, is provided in **Annex 5**.

The development of PGIs has transformed the genomic research landscape and holds potential for translation, most notably in healthcare.[52] For example, PGIs for diagnosable health conditions (such as heart disease) are **undergoing pilot trials** to assess their utility in informing population-level screening and management of health conditions. Nonetheless, a range of social and ethical issues pertain, as highlighted in a recent report – **"Predicting: the Future of Health?"** from the Ada Lovelace Institute and the Nuffield Council on Bioethics.

The remainder of this report focuses largely on the scientific insights and ethical implications of PGIs for social and behavioural phenotypes relating to education. These relate to both basic research on the genetic and environmental contributions to phenotypic variation across development (**Section 4**), and to some of the issues that may arise when considering translation into education contexts (**Section 5**). While this report does not aim to be exhaustive, we synthesise some of the key areas of contention identified during our research.

51 Choi SW, Mak TS, and O'Reilly PF (2020) Tutorial: a guide to performing polygenic risk score analyses **15(9)**: 2759–772.

52 Lennon NJ, Kottyan LC, Kachulis C, *et al* (2024) Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations *Nature Medicine* **30**: 480–7.

3 Genomic contributions to learning and education: key insights

Key messages

- In some **genomics** research, the word 'attainment' has been used to refer to the number of years spent in education, which is different from how the term is typically used in social science research and education.
- Number of years in education ('EduYears') is the most widely studied social **phenotype** in genomics, largely due to data availability.
- Like medical conditions, behavioural and social phenotypes related to education are highly **polygenic** and multifactorial.
- Specific genetic variants of small effect have been identified and tend to be **pleiotropic**, predicting multiple phenotypes.
- The EduYears **polygenic index** (PGI) predicts a larger share of individual educational differences than some other indicators (such as family socioeconomic status), but it is less predictive than prior achievement.
- While thousands of genetic associations have been identified, much of the polygenic component of social and behavioural phenotypes remains to be explored.
- Associations identified between social and behavioural phenotypes and genetic variants are correlational, and establishing causal paths from specific genetic variants to brain and behaviour is complex.
- The scientific and wider research community is starting to pay attention to wider discourse of the sociopolitical and ethical implications of genomics research.

At the time of writing, more than 6,000 **genome-wide association studies** (GWASs) have been performed across the medical, behavioural and social sciences for approximately 3,300 phenotypes. Early GWASs often failed to identify robust associations, as genetic **effect sizes** were much smaller than expected and sample sizes were not large enough to detect them. However, over the past five years –

Continued >>

thanks to expanded data-sharing practices and international collaborations – the average size of a GWAS has tripled to include around 140,000 participants. This has substantially increased the number of genetic associations identified.[53] A few GWASs are much larger than this; for instance, the latest GWAS for height comprised 5.4 million adults.[54]

In this section we begin by describing some of the most studied phenotypes that link to education and educational outcomes. We then provide an overview of findings from recent large-scale GWASs, including the generation of PGIs for educationally relevant phenotypes, and note some of the ethical concerns that have arisen as a consequence of these efforts.

3.1 Phenotypes that are relevant to education

For decades, twin and adoption study research has consistently demonstrated that nearly every aspect of individual differences in human behaviour and functioning are heritable.[55] In the context of education, this includes how many years one spends in education, performance in standardised tests ('achievement'), wellbeing, personality, attention, communication, behavioural needs, specific learning difficulties, socio-emotional phenotypes and even A-level subject choice.[56] This section of the report explores what is known about the genomic contributions to some of these heritable phenotypes. See **Annex 3** for more information about **heritability**.

The most extensively studied educational phenotype is the **number of years spent in formal education**, or 'EduYears'.[57] This measure is often referred to as 'attainment' in the genomics literature, which can be confusing as this is not how attainment is typically defined in the social sciences, education practice or policy.[58] Outside of genomics, attainment is an umbrella term that can also encompass performance in standardised assessments and tests or high-stakes exams. As a result, 'attainment' as it is often used in genomics research captures only one facet of education. Years

53 Abdellaoui A, Yengo L, Verweij KJH, and Visscher PM (2023) 15 years of GWAS discovery: realizing the promise *The American Journal of Human Genetics* **110(2)**: 179–94.

54 Yengo L, Vedantam S, Marouli E, *et al* (2022) A saturated map of common genetic variants associated with human height *Nature* **610**: 704–12.

55 Turkheimer, E (2000) Three laws of behavior genetics and what they mean *Current Directions in Psychological Science* **9(5)**: 160–4.

56 Rimfeld K, Ayorech Z, Dale P, *et al* (2016) Genetics affects choice of academic subjects as well as achievement *Scientific Reports* **6**: 26373.

57 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

58 Connelly R, Gayle V, and Lambert PS (2016) A review of educational attainment measures for social survey research *Methodological Innovations* **9**.

spent in education may correlate with variation in qualifications obtained, educational experiences and the nature of education undertaken, but it is not synonymous with these.

The primary reason that number of years spent in education is examined in genomics research is that it can be inferred from routinely collected demographic data, specifically 'highest level of education achieved'. This data can be reported directly by research participants in genomic cohorts and biobanks without requiring complex linkage to educational databases such as England's National Pupil Database (NPD). As a result, researchers are able to use this to calculate years-of-education equivalents across international cohorts, and conduct large-scale GWASs for 'EduYears'. To avoid confusion, we use the term 'years spent in education' (or 'EduYears') to be consistent with the actual measures used in these studies, rather than referring to it as attainment. Importantly, variation in years spent in education is recognised as a predictor of numerous life outcomes, such as **income**, health,[59] and even who you choose to partner with.[60]

Performance in standardised tests or exit exams across subjects have also been examined in GWAS frameworks, where they are referred to as 'educational achievement' or 'educational performance'.[61] Use of the term 'performance' in this way can be another source of confusion, because in UK education contexts, whilst performance usually relates to average exam grades, it may also take into account other indicators such as prior attainment and socioeconomic status – factors not routinely accounted for in genomic analyses.

Another phenotype examined is intelligence. Also called general cognitive ability, intelligence refers to a general mental capacity to reason, learn from experience, and understand complex ideas and concepts. This general factor, or 'g', is one of the most consistently documented findings in over a century of psychological research.[62] It reflects the observation that individuals who perform well on one type of cognitive test tend to do well on others.[63] Intelligence is a reliable predictor of educational,[64]

59 Balaj M, Henson CA, Aronsson A, *et al* (2024) Effects of education on adult mortality: a global systematic review and meta-analysis *The Lancet Public Health* **9(3)**: e155–65.

60 Mare, RD (1991) Five decades of educational assortative mating *American Sociological Review* **56(1)**: 15–32.

61 Rajagopal VM, Ganna A, Coleman JRI, *et al* (2023) Genome-wide association study of school grades identifies genetic overlap between language ability, psychopathology and creativity *Scientific Reports* **13**: 429.

62 Spearman C (1904) 'General intelligence', objectively determined and measured *American Journal of Psychology* **15**: 201–92.

63 Carroll JB (1993) *Human cognitive abilities: a survey of factor-analytic studies* (Cambridge: Cambridge University Press).

64 Deary IJ, Strand S, Smith P, and Fernandes C (2007) Intelligence and educational achievement *Intelligence* **35**: 13–21.

occupational[65] and health outcomes[66] and has been shown to remain relatively stable throughout adult life.[67]

However, this is not to suggest that intelligence is immutable. Intelligence scores are subject to change both within and between people; schooling has been shown to raise intelligence,[68] as has the rearing environment.[69] Discourse related to the measurement and interpretation of intelligence scores, such as its dependence on cultural context and socioeconomic status, is not covered in this report.[70]

Intelligence research continues to raise ethical concerns, especially when combined with genetics.[71] These concerns stem, in part, from historical abuse of intelligence testing and research findings. For instance, in the early 1900s cognitive screening in the US was used to justify the exclusion of individuals deemed intellectually 'inferior', particularly those from non-European countries.[72] Similarly, eugenic ideologies influenced policies such as forced sterilisation, targeting individuals with intellectual disabilities and other marginalised groups under the guise of preventing the transmission of 'undesirable' phenotypes.[73] Measures of intelligence are also the subject of continued ethical debate about their appropriateness and inclusivity.

This history and its legacy can make it challenging to justify genomic research into intelligence and other social and behavioural phenotypes, especially if there are minimal potential benefits to individuals and society – beyond the advancement of knowledge itself – alongside potential harms. The risks of genetic determinism, discrimination and stigmatisation persist. Moreover, evidence suggests that the findings from GWASs of intelligence and years spent in education are particularly susceptible to misuse and misappropriation. Specific examples include erroneous claims of a genetic basis for differences in intelligence between racial groups, and the justification of acts of violence.[74]

---

65 Strenze T (2007) Intelligence and socioeconomic success: a meta-analytic review of longitudinal research *Intelligence* **35**: 401–26.

66 Calvin CM, Batty GD, Der G, *et al* (2017) Childhood intelligence in relation to major causes of death in 68 year follow-up: prospective population study *British Medical Journal* **28(357)**: j2708.

67 Deary IJ, Pattie A, and Starr JM (2013) The stability of intelligence from age 11 to age 90 years: the Lothian birth cohort of 1921 *Psychological Science* **24**: 2361–8.

68 Brinch CN and Galloway TA (2012) Schooling in adolescence raises IQ scores *Proceedings in National Academy of Sciences* U.S.A. **109(2)**: 425–30.

69 Willoughby EA, McGue M, Iacono WG, and Lee JJ (2021) Genetic and environmental contributions to IQ in adoptive and biological families with 30-year-old offspring *Intelligence* **88**: 101579.

70 Lubinski D (2025) Education, intelligence, placement, and selection: a discussion of paradoxes and fairness *Intelligence* **108**: 101881.

71 Hayden EC (2013) Ethics: taboo genetics *Nature* **502**: 26–8.

72 Blinkhorn S (2019) Early US immigrants were tested for cognitive impairment, not IQ *Nature* **574(7776)**: 36.

73 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

74 Ibid.

Large genomic cohorts and biobanks contain information on many other phenotypes and conditions. This has led to statistically well-powered GWASs being conducted for a number of heritable phenotypes that relate to, or may be of interest to, education. We include some examples of these.

3.2 Polygenicity: genetic associations of very small effect have been identified

As mentioned above, the widespread availability of 'EduYears' data for cohort studies, coupled with the creation of the **Social Science Genetic Association Consortium** (SSGAC) for data sharing, has made years spent in education one of the most extensively studied social phenotypes in genomics. To date, four large-scale GWASs have been performed by the SSGAC for this measure. The first, published in 2013, involved 126,559 participants, making it one of the largest GWASs of 'EduYears' ever performed at the time.[75] As sample sizes increased, so did the number of genetic variants identified. The fourth and most recent SSGAC GWAS included around 3 million adults and identified 3,952 common genetic variants associated with years spent in education. A consistent finding across all four GWASs is that each identified genetic variant accounts for only a tiny fraction (at most 0.02%) of differences between individuals in years spent in education.[76]

A comprehensive summary of GWAS findings for other behavioural phenotypes related to education is beyond the scope of this report. However, several noteworthy studies – in terms of size – are described here.

The most recent GWAS for intelligence comprised 269,867 adults from 11 European cohorts and identified 205 genetic associations.[77] This replicated many of the associations found in earlier GWASs of general cognitive function.[78] A 2022 GWAS of dyslexia based on more than 50,000 self-reported dyslexic adults and around 1 million controls, identified 42 genetic associations – many of which were novel.[79] The largest GWAS for attention deficit hyperactivity disorder (ADHD) analysed 38,691

75 Rietveld CA, Medland SE, Derringer J, *et al* (2013) GWAS of 126,559 individuals identifies genetic variants associated with educational attainment *Science* **340(6139)**: 1467–71.

76 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

77 Savage JE, Jansen PR, Stringer S, *et al* (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence *Nature Genetics* **50(7)**: 912–19.

78 Sniekers S, Stringer S, Watanabe K, *et al* (2017) Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence *Nature Genetics* **49**: 1107–12.

79 Doust C, Fontanillas P, Eising E, *et al* (2022) Discovery of 42 genome-wide significant loci associated with dyslexia *Nature Genetics* **54**: 1621–29.

individuals with ADHD and 186,843 controls, and identified 27 genomic regions associated with ADHD.[80]

Other published GWASs of interest to educators and educational researchers include those for exit exam performance in languages and mathematics derived from the Danish education register;[81] reading and language skills;[82] crystallised and fluid cognitive abilities derived from online testing;[83] executive function;[84] non-cognitive skills;[85] autism;[86] personality;[87] and mental health disorders such as depression and anxiety.[88] Online GWAS catalogues such as the **EMBL-EBI Catalog** and the **IEU Open GWAS project** can be queried to identify whether a GWAS for a specific phenotype has been published, the results of the study, and whether the **GWAS summary data** are available to download.

A key finding to emerge from these large-scale GWASs is that complex phenotypes related to education are highly **polygenic**. This is not a finding unique to phenotypes relevant to education: complex medical conditions (such as cardiovascular disease) and physical phenotypes (such as height and weight) are also influenced by thousands of genetic variants, each of tiny (but detectable) effect, along with environmental factors.

80 Demontis D, Walters GB, Athanasiadis G, *et al* (2023) Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains *Nature Genetics* **55**: 198–208.

81 Rajagopal VM, Ganna A, Coleman JRI, *et al* (2023) Genome-wide association study of school grades identifies genetic overlap between language ability, psychopathology and creativity *Scientific Reports* **13**: 429.

82 Eising E, Mirza-Schreiber N, de Zeeuw EL, *et al* (2022) Genome-wide analyses of individual differences in quantitatively assessed reading- and language-related skills in up to 34,000 people *Proceedings of National Academy of Sciences* U.S.A. **119(35)**: e2202764119.

83 Carey C, Huang Y, Strong RW, *et al* (2021) Shared and distinct genetic influences between cognitive domains and psychiatric disorder risk based on genome-wide data *Biological Psychiatry* **89(9)**.

84 Hatoum AS, Morrison CL, Mitchell EC, *et al* (2023) Genome-wide association study shows that executive functioning is influenced by GABAergic processes and is a neurocognitive genetic correlate of psychiatric disorders *Biological Psychiatry* **93(1)**: 59–70.

85 Demange PA, Malanchini M, Mallard TT, *et al* (2021) Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction *Nature Genetics* **53**: 35–44.

86 Grove J, Ripke S, Als TD, *et al* (2019) Identification of common genetic risk variants for autism spectrum disorder *Nature Genetics* **51**: 431–44.

87 Gupta P, Galimberti M, Liu Y, *et al* (2024) A genome-wide investigation into the underlying genetic architecture of personality traits and overlap with psychopathology *Nature Human Behaviour* **8**: 2235–49.

88 Sullivan PF, Agrawal A, Bulik CM, *et al* (2018) Psychiatric genomics: an update and an agenda *American Journal of Psychiatry* **175(1)**: 15–27.

3.3 Measurement of phenotypes relevant to education

The large sample size required for GWASs means that the highly polygenic phenotypes being examined are often measured in a very limited way. For instance, as discussed above, the GWASs of years spent in education were performed because more detailed measures of educational attainment and achievement were not available (and still are not available) in the large genomic cohorts needed for a GWAS. In essence, there is a trade-off between sample size and measurement, and in most instances sample size wins, leading to a relatively crude measure of the phenotype of interest or proxies, such as EduYears, being used. It is worth noting that the outcome measures that educators and policymakers are particularly concerned with are overall attainment levels, and gaps between groups. These data are available in the NPD, but in most instances genomic cohorts are not linked to the NPD. This divergence between what can be measured in genomic studies and what stakeholders prioritise highlights an important challenge in aligning genomic research with real-world educational goals.

The dyslexia GWAS described above is another example of this trade-off; it used data collected from the commercial direct-to-consumer (DTC) company 23&Me that relied on a single participant self-report item: the response to the question 'Have you been diagnosed with dyslexia?'.[89] More detailed and well-validated assessments of reading difficulty are available, but as these require resource-intensive administration by educational psychologists or appropriately trained researchers, they are not routinely collected for very large research cohorts or biobanks. The impact of taking a minimal approach to measurement of the phenotype being studied in a GWAS is only starting to be understood.[90]

When more detailed and rigorous measurements are collected across smaller research cohorts, meta-analytical approaches can be employed to combine data from multiple independent cohorts. This has been facilitated by the formation of international consortia that co-ordinate data sharing and access across research groups, such as the **SSGAC** (for years spent in education), the **Psychiatric Genomics Consortium** (for psychiatric conditions) and the **GenLang** consortium (for speech, reading, language and related skills).

However, meta-analyses can be complicated by diversity in how phenotypes are measured across cohorts, as well as variability arising from differences in participant recruitment criteria, age, time period, and environmental or social contexts. Many of the same data sources are included in multiple consortia, meaning that the same populations – predominantly those of European **genetic ancestry** – are analysed repeatedly.

89 Doust C, Fontanillas P, Eising E, *et al* (2022) Discovery of 42 genome-wide significant loci associated with dyslexia *Nature Genetics* **54**: 1621–9.

90 Cai N, Revez JA, Adams MJ, *et al* (2020) Minimal phenotyping yields genome-wide association signals of low specificity for major depression *Nature Genetics* **52**: 437–47.

3.4 Missing heritability: much of the polygenic component remains to be discovered

While many genetic variants have been identified by GWASs, if the effects of each are added up, they account for only a small fraction of the phenotype's **heritability** as estimated using **twin studies**. This has been called the '**missing heritability** problem'.[91] However, this gap is much smaller if SNP-based heritability is considered, i.e. if the effects of all genetic variants examined in a GWAS are taken into account (see **Annex 3**).

A key finding to emerge is that the SNP-based heritability of behavioural phenotypes related to education is larger than the variance explained by the set of genetic associations identified in a GWAS, but smaller than twin-based heritability values. For example, when summed together, the genetic variants associated with years spent in education account for about 13% of differences between individuals, while twin studies estimate that about 40% of these differences are due to genetic variation.[92] The SNP heritability for years in education is about 20%, between these two values.[93] A similar pattern is observed for intelligence (see Figure 2).[94] For phenotypes such as ADHD and anxiety, the SNP heritability is generally less than half the twin-based heritability, averaging at around 37%.[95]

Figure 2: Missing heritability for intelligence

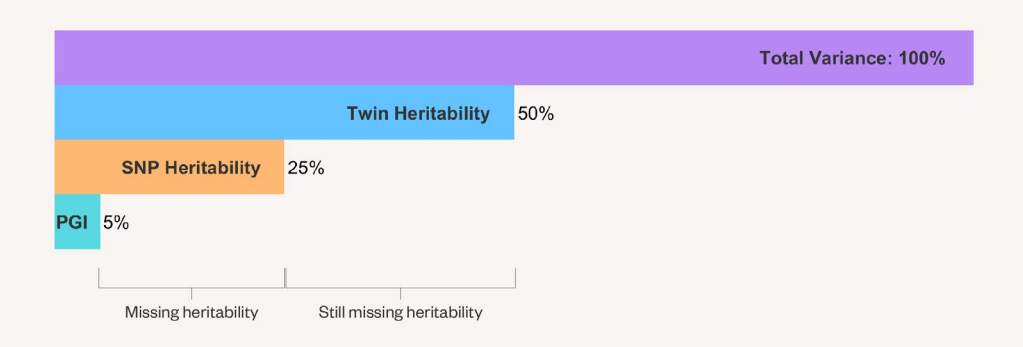


Figure 2 shows heritability estimates for intelligence based on the intelligence PGI, all SNPs measured in the GWAS (SNP heritability) and twin studies (twin heritability). Two types of 'missing heritability' have been identified for behavioural and social phenotypes, including intelligence and years in education. A gap exists between all genetic associations identified in a GWAS (~5% for intelligence) and SNP-based

91 Manolio T, Collins F, Cox N, *et al* (2020) Finding the missing heritability of complex diseases *Nature* **461**: 747–53.

92 Silventoinen K, Jelenkovic A, Sund R, *et al* (2020) Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts *Scientific Reports* **10**: 1–11.

93 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

94 Plomin R and von Stumm S (2018) The new genetics of intelligence *Nature Reviews Genetics* **19(3)**: 148–59.

95 Plomin R (2022) The next 10 years of behavioural genomic research *JCPP Advances* **2(4)**: e12112.

heritability (~25%), which can be narrowed by increasing GWAS sample sizes. A gap also exists between SNP heritability (~25%) and twin heritability (~50%), the narrowing of which will require alternative approaches that consider rare variants and interactions between genes and environments.

This indicates two things. Firstly, very weak-effect common genetic variants are being missed by the discovery GWASs due to inadequate **statistical power**. Increasing sample sizes even further is a way around this issue, and is a key reason why GWASs continue to be conducted for years spent in education using ever larger numbers of participants. Secondly, if twin-based estimates are broadly correct,[96] the gap between SNP and twin-based heritability indicates that other sources of genetic influence – beyond the additive effects of common variants – are important. The sources of this 'still-missing' heritability are debated, but likely involve the contribution of low-frequency (rare) variants of larger effect, and gene–gene and **gene–environment interaction** effects.[97]

Like **polygenicity**, the gap between variance accounted for by identified genetic associations and SNP-based and twin-based heritability estimates is not unique to behavioural and social phenotypes, as it is also observed for medical conditions and physical phenotypes.[98]

3.5 Pleiotropy: genetic variants associate with multiple phenotypes

An early observation from GWASs was that genetic associations identified for one social or behavioural phenotype often also associate with multiple other phenotypes. For instance, genetic variants associated with ADHD also overlap with those linked to schizophrenia, major depressive disorder and autism.[99] This contrasts with neurological conditions such as Parkinson's and Alzheimer's disease, which appear to be more genetically distinct.[100]

Rather than focusing on one genetic association at a time, newly developed methods (see **Annex 3**) allow researchers to assess the extent to which all genomic variation measured in a GWAS overlaps between pairs of phenotypes – a concept called **genetic correlation**. These studies have provided evidence of widespread **pleiotropy**,

96 Wolfram T and Morris D (2023) Conventional twin studies overestimate the environmental differences between families relevant to educational attainment *NPJ Science of Learning* **8**: 24.

97 Brandes N, Weissbrod O, and Linial M (2022) Open problems in human trait genetics *Genome Biology* **23**: 131.

98 Ibid.

99 Demontis D, Walters GB, Athanasiadis G, *et al* (2023) Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains *Nature Genetics* **55**: 198–208.

100 Brainstorm Consortium (2018) Analysis of shared heritability in common disorders of the brain *Science* **360(6395)**: eaap8757.

where the same genetic variants affect multiple phenotypes across domains such as years in education, health, neurodevelopment and psychiatry.[101] Some of these genomic relationships are unexpected, such as the one reported between dyslexia and pain.[102]

From a research standpoint, evidence of pleiotropy can strengthen observational longitudinal studies by helping researchers to develop hypotheses about the potential reasons for genetic correlation. For example, does the moderate positive genetic correlation between dyslexia and ADHD arise from shared biological mechanisms, with genetic variants influencing a common cognitive process or property underlying both (e.g. attention or synaptic pruning)? Or does it arise due to the causal effect of **genetic variation** influencing ADHD via its effect on dyslexia? Emerging approaches that utilise genomic and GWAS data may help researchers start to tackle these types of questions.[103]

More directly, pleiotropy means that an individual's PGI (**Section 2.10**) for one phenotype can also predict other phenotypes. For example, genetic variants associated with higher intelligence are also associated with a higher risk of anorexia nervosa,[104] while genetic variants associated with more years in education link to increased risk for schizophrenia.[105] This has ethical implications for the interpretation and potential application of PGIs in settings such as education (see **Section 5.2**), as an individual's PGI for one phenotype may inadvertently reveal information about other unrelated or sensitive phenotypes. In such a scenario, how should PGI information be responsibly and clearly communicated, informed consent managed, and privacy protected?

3.6 Understanding the biology has proved difficult

As social and behavioural phenotypes related to education are highly **pleiotropic** and **polygenic**, biological interpretation of GWAS results has proved difficult. As discussed in **Section 2.8**, **gene–environment correlation** complicates matters further, as it means that polygenic influences can manifest not only via biological pathways but also through social or environmental mechanisms.

101 Bulik-Sullivan BK, Po-Ru L, Finucane HK, *et al* (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies *Nature Genetics* **47(3)**: 291–5.

102 Doust C, Fontanillas P, Eising E, *et al* (2022) Discovery of 42 genome-wide significant loci associated with dyslexia *Nature Genetics* **54**: 1621–29.

103 Frei O, Holland D, Smeland OB, *et al* (2019) Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation Nature Communications 10: 2417; Grotzinger AD, Rhemtulla M, de Vlaming R, *et al* (2019) Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits *Nature Human Behaviour* **3**: 513–25.

104 Hill WD, Harris SE, and Deary IJ (2019) What genome-wide association studies reveal about the association between intelligence and mental health *Current Opinions in Psychology* **27**: 25–30.

105 Lee JJ, Wedow R, Okbay A, *et al* (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals *Nature Genetics* **50(8)**: 1112–21.

It can be hard to reconcile how DNA differences – which must act through biological mechanisms such as altering gene expression or neural development – can be mediated by the environment. A hypothetical scenario often used to illustrate socially mediated genetic effects involves discrimination based on a genetically influenced phenotype.

For instance, imagine a society where blue-eyed children are systematically disadvantaged in education (e.g. given fewer learning resources and support), while brown-eyed children are favoured. In this scenario, the genetic variants for eye colour influence educational outcomes indirectly by eliciting differential treatment from others within the social environment. This produces a link between genetic variants for eye colour and educational outcomes through **gene–environment correlation**. If a GWAS were conducted in this society and the results were used to create a PGI for educational attainment, children with blue eyes would likely have a lower educational attainment PGI than children with brown eyes. In this scenario, the genetic associations – and by extension, the PGI – are indexing aspects of the child's social environment that contribute to disparities in education.

There is now strong evidence that many of the genetic associations for years spent in education not only reflect the biology of the individual, but also capture aspects of the individual's demographic characteristics and social environment that influence educational outcomes (see **Section 4.4**).

Unravelling the complex biological, social and environmental paths linking highly polygenic and pleiotropic genomic influences on social and behavioural phenotypes related to education remains a challenge for researchers and will require other analytical approaches, including biological and psychological developmental models.

For now, insights into specific biological mechanisms underpinning differences in learning remain scarce. For example, while the dyslexia GWAS implicated brain-expressed genes, these genes do not readily map to biological pathways or cognitive models of reading.[106] Similarly, the genetic variants identified in the SSGAC's series of EduYears GWASs are located close to genes that are expressed during neurodevelopment, but do not include those involved in glial cell function. The authors note that this was unexpected, given neuroscientific evidence for the role of glial cells in learning and memory.[107] Finally, the most recent GWAS of intelligence pointed to contributions from general properties of the brain including neurogenesis (the number of neurons produced), neuronal differentiation (the specialisation of neurons) and synaptic communication.[108]

106 Doust C, Fontanillas P, Eising E, *et al* (2022) Discovery of 42 genome-wide significant loci associated with dyslexia *Nature Genetics* **54**: 1621–29.

107 Lee JJ, Wedow R, Okbay A, *et al* (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals *Nature Genetics* **50(8)**: 1112–21.

108 Savage JE, Jansen PR, Stringer S, *et al* (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence *Nature Genetics* **50(7)**: 912–19.

3.7 Polygenic indices: genomic predictors of social and behavioural phenotypes

Even if researchers do not fully understand (or cannot explain) the causal biological mechanisms and social processes that link DNA variation to specific phenotypes, this lack of 'explanation' does not prevent use of the variants for predictive purposes.[109] A useful analogy is smoking: the correlation between smoking and lung cancer was enough to drive successful public health interventions long before the biological pathways of carcinogenic effects were fully understood. Similarly, as described in **Section 2.10**, the genetic variants identified in a GWAS can be summed together to create a PGI that – like the smoking/lung cancer example – does not require knowledge of the underlying causal processes but still predicts population-level variation in a phenotype (see **Annex 5**).

Over the past five years, the predictive power of PGIs for social and behavioural phenotypes has steadily increased as GWAS sample sizes have grown. This is neatly illustrated by the PGI for years spent in education: at the population level, the PGI calculated from the SSGAC's 2018 GWAS of around 1.1 million individuals predicts 11–13% of variation in 'EduYears' in independent samples.[110] The 2022 PGI of around 3 million participants predicts 12–16% of variation in 'EduYears'.[111] This percentage makes the 'EduYears' PGI one of the most predictive PGIs in the behavioural sciences. Accounting for 12–16% of the variance makes it a stronger predictor of years spent in education than measures of household income and marital status, and of similar strength to parental educational attainment or performance on cognitive tests.[112]

PGIs are now available for other phenotypes related to education. For example, cognitive PGIs have been published that predict 6% of variation in reading skills,[113] 5% in general cognitive function (intelligence[114]), and 15% in tested school performance at age 16.[115] While not a direct measure of cognition, the EduYears PGI accounts for more variance in intelligence than the intelligence PGI, explaining 10% of individual differences.[116]

109 Plomin R and von Stumm S (2022) Polygenic scores: prediction versus explanation *Molecular Psychiatry* **27**: 49–52.

110 Lee JJ, Wedow R, Okbay A, *et al* (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals *Nature Genetics* **50(8)**: 1112–21.

111 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

112 Ibid.

113 Doust C, Fontanillas P, Eising E, *et al* (2022) Discovery of 42 genome-wide significant loci associated with dyslexia *Nature Genetics* **54**: 1621–29.

114 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49; Savage JE, Jansen PR, Stringer S, *et al* (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence *Nature Genetics* **50(7)**: 912–19.

115 Allegrini AG, Selzam S, Rimfeld K, von Stumm S, Pingault JB, and Plomin R (2019) Genomic prediction of cognitive traits in childhood and adolescence *Molecular Psychiatry* **24(6)**: 819–27.

116 Plomin R and von Stumm S (2018) The new genetics of intelligence *Nature Reviews Genetics* **19(3)**: 148–59.

For other phenotypes such as anxiety, depression and ADHD, PGIs predict an average of 4% of the variance.[117] Finally, a PGI for externalising behaviours (e.g. hyperactivity, impulsivity, aggression) was recently published that explains 10% of variance, which the authors note is similar to other predictor variables including family income and neighbourhood disadvantage.[118]

The key takeaway is that some PGIs are approaching the predictive power of demographic measures routinely used in the social sciences.

The genomics research community generally seem confident that the proportion of population-level variance explained by PGIs will increase in the coming years as GWAS samples become larger; finer-grained measures of educational phenotypes become available (i.e. individuals are tested rather than relying on self-reported items, and more cohorts are linked to educational databases); rarer DNA variants are assessed;[119] analytical strategies improve;[120] and multiple PGIs for genetically correlated phenotypes are used.[121] While there is a lack of agreement about how predictive PGIs will become, and within what time frame, it is generally assumed that the ceiling for PGI-based prediction will be no greater than the heritability estimates from twin studies.

3.8 Ethically responsible scientific conduct and communication

As described in **Section 2.6**, GWAS researchers typically group participants into genetically similar populations, referred to as **genetic ancestry**, with the vast majority of educationally relevant GWASs having been conducted in populations of European genetic ancestry.[122] Further, the results of a GWAS analysis do not 'port' well across populations. The combination of this bias and lack of portability means that PGIs for many phenotypes related to education are less predictive for individuals of non-European genetic ancestry.[123] For example, the EduYears PGI derived from the SSGAC's 2022 GWAS of a European genetic ancestry sample predicts 12–16% of

117 Ibid.

118 Karlsson Linnér R, Mallard TT, Barr PB, *et al.* (2021) Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction *Nature Neuroscience* **24**: 1367–76.

119 Wainschtein P, Jain D, Zheng Z, *et al* (2022) Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data *Nature Genetics* **54(3)**: 263–73.

120 Márquez-Luna C, Gazal S, Loh PR, *et al* (2021) Incorporating functional priors improves polygenic prediction accuracy in UK Biobank and 23andMe data sets *Nature Communications* **12**: 6052.

121 Procopio F, Liao W, Rimfeld K, *et al* (2024) Multi-polygenic score prediction of mathematics, reading, and language abilities independent of general cognitive ability *Molecular Psychiatry* 31 July.

122 Mills MC and Rahal CA (2019) A scientometric review of genome-wide association studies *Communications in Biology* **2**: 9.

123 Duncan L, Shen H, Gelaye B, *et al* (2019) Analysis of polygenic risk score usage and performance in diverse human populations *Nature Communications* **10**: 3328.

the variance in European populations but only about 2% in a sample of African genetic ancestry.[124] There have been no massive-scale GWASs for years spent in education reported for more diverse populations.

Because genomic effects are context-dependent, PGI predictions can also vary when the environmental conditions of a population differ from those of the discovery GWAS from which the PGI was derived. For example, even within the same population, the performance of the EduYears PGI can differ based on social and environmental characteristics such as age, sex and socioeconomic status.[125] While this variability is not unexpected, it underscores the complexity of PGIs and presents a challenge to the equitable translation of educational genomic research findings, mirroring similar challenges observed in healthcare.[126]

Another concern arises when the results of genetic ancestry analyses performed in a GWAS are misappropriated or misinterpreted to imply a genetic and biological basis to race.[127] This becomes especially problematic when combined with GWAS findings to suggest evidence of a biological explanation for mean differences in phenotype between racial groups.[128] Such misinterpretations risk perpetuating harmful and scientifically unfounded notions of inherent racial group differences, and detract from addressing the true underlying drivers of these disparities – namely, social, environmental and historical inequities.

In response to these risks, the scientific community is actively working to develop ethical guidance for best practice for GWAS researchers, and frameworks to enable and support responsible communication of GWAS findings to the public.[129] While a detailed review of emerging recommendations is beyond the scope of this report, we summarise key points raised in conversations with experts and in the literature.

The National Academies of Science, Engineering, and Medicine (NASEM) have produced evidence-based guidelines on the appropriate use of population descriptors such as race, ethnicity and ancestry in genomics research.[130]

---

124 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

125 Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, and Przeworski M (2020) Variable prediction accuracy of polygenic scores within an ancestry group Elife **9**: e48376.

126 Farmer H (2024) Predicting: the future of health? (London: Ada Lovelace Institute).

127 Wills M (2017) Are clusters races? A discussion of the rhetorical appropriation of Rosenberg *et al*.'s 'Genetic structure of human populations' *Philosophy, Theory, and Practice in Biology* **9(12)**.

128 Duncan RG, Krishnamoorthy R, Harms U, *et al* (2024) The sociopolitical in human genetics education *Science* **383(6685)**: 826–8.

129 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41; Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–49.

130 National Academies of Sciences, Engineering, and Medicine (2023) *Using population descriptors in genetics and genomics research: a new framework for an evolving field* (Washington, DC: National Academies Press).

They recommended that researchers provide clear definitions and justification for such terms in their studies, using genetic ancestry only when it is necessary to answer the research question. Their recommendations have already informed new publication standards for biomedical research journals.[131] These efforts represent a step towards promoting more responsible research and its communication; reducing harmful and inaccurate typological thinking about individuals; and ensuring that all populations benefit from genomic advances.

A practical set of steps relates to how researchers communicate their research findings to a lay audience, and how to better engage in the sociopolitical context of genomics research.[132] For example, providing a **frequently asked questions (FAQ)** page alongside a research article offers an opportunity to explain the context, scope and limitations of the research to a lay audience. This can be paired with a disclaimer within the research article that explicitly states how the results should not be used or interpreted.[133] Together, these strategies may enhance public understanding and reduce the risk of discrimination, stigmatisation and labelling of individuals and groups. It can also help to discourage premature commercial, practice-based and policy applications of research findings. The SSGAC serves as an exemplar in this regard,[134] and a publicly available FAQ repository for social and behavioural genomic research has recently been launched.[135] The issue of responsible communication raises the broader question of whether scientific responsibility ends with publication of a study. Should efforts be made to track where the study is being discussed or misappropriated,[136] and should such instances be responded to, for example, by writing an opinion piece or engaging with the media? While such efforts may be difficult for individual research groups to co-ordinate, they could be supported by dedicated teams at the funding, institutional, or departmental levels.[137]

A third recommendation is for the meaningful involvement of the public and communities in research about them and directly affecting them. In the cognitive sciences, public involvement may help to identify ethical concerns; lower barriers to participation; ensure research is relevant and findings are accessible; and

131 Feero WG, Steiner RD, Slavotinek A, *et al* (2024) Guidance on use of race, ethnicity, and geographic origin as proxies for genetic ancestry groups in biomedical publications *JAMA* **331(15)**: 1276–8.

132 Duncan RG, Krishnamoorthy R, Harms U, *et al* (2024) The sociopolitical in human genetics education *Science* **383(6685)**: 826–8.

133 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

134 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

135 Martschenko DO, Domingue BW, Matthews LJ, and Trejo S (2021) FoGS provides a public FAQ repository for social and behavioral genomic discoveries *Nature Genetics* **53**: 1272–4.

136 Carlson J and Harris K (2020) Quantifying and contextualizing the impact of bioRxiv preprints through automated social media audience segmentation *PLoS Biology* **18(9)**: e3000860.

137 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

support the diversification of data collection efforts.[138] A highly collaborative form of public involvement is co-production research, which actively engages individuals, families and communities in designing and conducting research studies. This can help to ensure a diverse and inclusive range of perspectives, while keeping the research both relevant and practically applicable in real-world settings.[139]

At least in the UK, there also appears to be a gap in compulsory training of genomics researchers in the societal and ethical implications of their work. This is not insurmountable: **online courses** and **resources**, including those collated by the **ELSI hub**, can provide researchers with some of the ethical tools they need to better evaluate and navigate the impact of their research.[140] This could be further strengthened by collaboration between geneticists and bioethical experts across all stages of a genomic study.[141]

Ensuring that genomic research on educationally relevant outcomes and phenotypes is conducted responsibly and communicated accurately is essential, not only for advancing knowledge but also for safeguarding against potential harms. Adoption of the steps described can help to minimise potential harms to individuals, groups and society, while advancing knowledge of the genetic contributions to variation in educationally relevant outcomes.

138 Garcini LM, Arredondo MM, Berry O, *et al* (2022) Increasing diversity in developmental cognitive neuroscience: a roadmap for increasing representation in pediatric neuroimaging research *Developmental Cognitive Neuroscience* **58**: 101167.

139 Redman S, Greenhalgh T, Adedokun L, *et al* (2021) Co-production of knowledge: the future *BMJ* **372**: n434.

140 Patel RA, Ungar RA, Pyke AL, *et al* (2024) Equity in science requires better ethics training: a course by trainees, for trainees *Cell Genomics* **4(5)**: 100554.

141 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

4 Challenges and opportunities in translating genomic insights to education

Key messages

- The predictive accuracy and portability of **polygenic indices** (PGIs) for educationally relevant **phenotypes** is limited, which constrains their potential use in educational settings.
- The implementation of PGIs for individual prediction raises significant ethical concerns, including risks of inequity, discrimination, stigmatisation and genetic determinism.
- The real-world phenomenon of **gene–environment correlation** poses a challenge to the interpretation of both environmental and genomic research.
- Genomic data are increasingly available from large-scale cohorts and are being harnessed to untangle the genetic and environmental contributions to phenotypes related to education.
- PGIs may offer a way to partially control for genetic **confounding** when studying environmental influences on educational outcomes.
- Integrating PGI data in family designs can help to build understanding of the intergenerational transmission of genetic predispositions to educationally relevant phenotypes.
- **Gene–environment interaction** research using PGIs may offer insights into the environments, experiences and interventions that are most effective for different children.
- **Mendelian randomisation** may help researchers to draw more reliable conclusions about cause and effect by using genomic data that are more robust to confounding and reverse causation.
- PGIs come with limitations, particularly when applied to individuals, so their use and interpretation across contexts requires careful consideration and expertise.

Continued >>

Some of the observed differences among individuals in educational, cognitive and behavioural phenotypes can be attributed to DNA variation. In **Section 2** we described the methodological approaches to discovering phenotype-associated DNA variants, and in **Section 3** we summarised what is known – and not known – about these contributions. We also identified some of the ethical issues that result from scientific limitations, misunderstanding and misinterpretation of knowledge, and how concerns might be minimised.

The first part of this section explores the scientific potential, and some of the ethical implications of, using PGIs for individual-level prediction of educationally relevant phenotypes. This application has generated tension and debate in the scientific community, with many researchers expressing concern that use of PGIs in their current form for such a purpose is both scientifically and ethically problematic.

In the second part, we examine some of the ways in which PGIs are currently being applied in the social and behavioural sciences as research variables as a means of improving understanding of environmental effects on educationally relevant phenotypes. While these basic research applications present new methodological opportunities, they, too, are debated, most notably due to the limitations and complexities of PGIs.

Gene-discovery research into educationally relevant phenotypes is a dynamic field. GWASs continue to be conducted on increasingly larger samples, with a growing emphasis on the inclusion of more genetically and environmentally diverse populations (see **Section 2.6**). The majority of GWAS research so far has been conducted using adult participants, but efforts are now underway to identify common genetic variants associated with phenotypes in infancy and childhood.[142] Discovery research efforts are also being expanded to consider a wider range of **genetic variation**, including the role of rare coding variants in cognitive function.[143] Finally, many studies are being published that examine how PGIs for a specific phenotype (such as ADHD) associate with related phenotypes (such as education and mental health) in independent clinical and population cohorts.[144] The pace of this research underscores the need for continued evaluation and consideration of both scientific and ethical implications to ensure that the knowledge gained is used responsibly and equitably.

142 Ronald A and Gui A (2024) The potential and translational application of infant genetic research *Nature Genetics* **56(7)**: 1346–54.

143 Chen CY, Tian R, Ge T, *et al* (2023) The impact of rare protein coding genetic variation on adult cognitive function *Nature Genetics* **55**: 927–38.

144 Ronald A, de Bode N, Polderman TJC (2021) Systematic review: how the attention-deficit/hyperactivity disorder polygenic risk score adds to our understanding of ADHD and associated traits *Journal of the American Academy of Child and Adolescent Psychiatry* **60(10)**: 1234–77.

4.1 Polygenic indices as tools for education?

As discussed in **Section 3**, PGIs have been developed for a range of behavioural and cognitive phenotypes and social measures – most notably years spent in education – that account for some of the observed differences between individuals in a population. A few researchers[145] have suggested that as a result of this predictive power, PGIs hold potential for individual prediction in educational settings, though this is not a universally held view.[146] This potential is based on the observation that PGIs can predict mean differences in phenotypes at the extremes. For example, Okbay *et al* (2022) grouped individuals based on their EduYears PGI, from low to high. They found that on average across two independent cohorts, 62% of individuals in the highest group (top 10% PGI) had a university degree, compared with 7% of individuals in the lowest group (bottom 10% PGI). A similar pattern is found for the intelligence PGI, where the mean IQ was 92 in the lowest PGI group and 108 in the highest.[147]

As a result of these observed differences at the PGI extremes, it has been suggested that PGIs might feasibly be used as indicators of an individual's genomic predisposition to specific phenotypes, in order to personalise education and learning. Educational curricula, learning environments or teaching styles could be tailored to an individual based on their PGI, with a focus on identifying and supporting those who are more likely to experience difficulties in school. However, as already discussed, a critical caveat is that these findings may not generalise to everyone or to all educational contexts, as the populations on which they are based are genetically and socioeconomically homogeneous.

145 Plomin R and von Stumm S (2018) The new genetics of intelligence *Nature Reviews Genetics* **19(3)**: 148–59; Plomin R (2018) *Blueprint: how DNA makes us who we are* (Cambridge, MA: The MIT Press).

146 Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–49.

147 Plomin R and von Stumm S (2018) The new genetics of intelligence *Nature Reviews Genetics* **19(3)**: 148–59.

Figure 3: Scatterplot of the relationship between the EduYears PGI derived from Okbay et al's 2022 GWAS of ~3 million individuals and reported years spent in education for two independent cohorts, the National Longitudinal Study of Adolescent to Adult Health (Add Health) and the Health and Retirement Study (HRS).

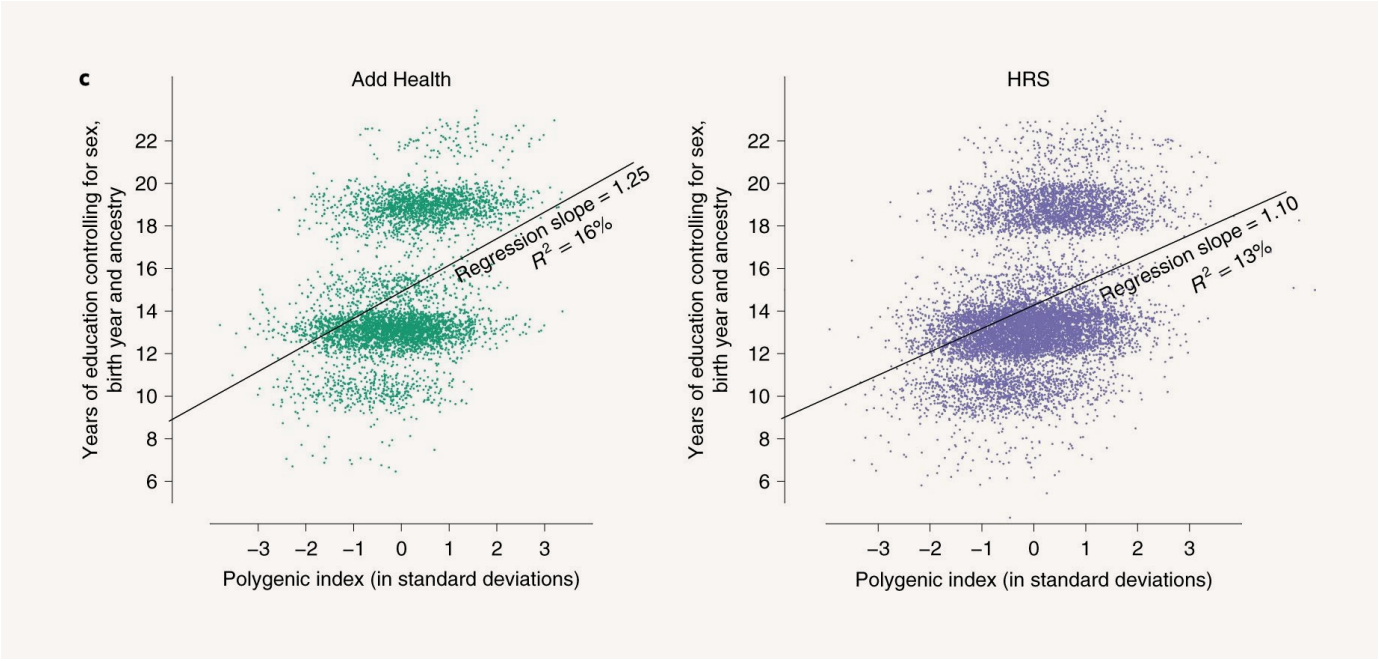


Figure 3 shows the relationship between the EduYears PGI (x-axis) and measured years in education (y-axis), with each point representing an individual. While there is an overall positive correlation between the PGI and years in education, there is a large amount of variability, even at the extremes of PGI values.

While PGIs can explain a chunk of population variance, they – like all predictors – are not accurate at the individual level. As a result, researchers have stated that PGIs for social and behavioural phenotypes are limited as tools for individual prediction.[148] The **FAQs accompanying** the SSGAC's latest GWAS of years spent in education clearly states this, warning that "it is important that participants/users understand that these individual results are not meaningful predictions and should be regarded essentially as entertainment. Failure to make this point clear risks sowing confusion and undermining trust in genetics research."

148 Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–49.

This warning was given because the average differences observed between extremes of PGIs mask a wide range of individual differences. For example, **Figure 3** illustrates that many individuals with a very high EduYears PGI spent fewer years in education than individuals with a very low PGI, and vice versa. The lack of predictive accuracy is because the PGI does not capture all the genomic influence on years spent in education. Even if it did, **heritability** for this measure is not 100%, and environmental context may modify or override PGI influences.

The best available predictor of an individual's educational achievement is prior attainment, measures of which are routinely collected as part of standard classroom practice or through standardised tests and public examinations, accounting for a high level of variance in later achievement.[149] Unlike PGIs, however, prior achievement data require a history of schooling and, as such, are not available at birth. By contrast, DNA differences are fixed and can be used to make predictions – in theory – from the moment of conception, or before schooling begins.[150]

How well does the EduYears PGI predict realised educational achievement, and how does it compare to other student background data readily available to schools, such as student age and sex[151], eligibility for free school meals, special educational needs status, and parental social background? Morris *et al* (2020) examined these questions in a UK cohort using the SSGAC's 2018 EduYears PGI. They found that while children with a higher PGI, on average, had higher exam scores than those with a lower PGI, at the individual level, the PGI was less accurate in predicting achievement than parental years of education and socioeconomic position.[152] However, parental background information is not always available, and parental predictors cannot distinguish between children within the same family. In contrast, PGIs are specific to an individual and, as such, can differentiate children within the same family.

Education already provides real-world examples where imperfect and 'noisy' predictors of individual achievement are problematic.[153] For instance, in England, students' prior achievement at Key Stage 2 is sometimes used to set target grades for GCSE subjects. The accuracy of these targets can vary, yet student, parent and teacher expectations often fail to account for this uncertainty.[154]

149 Rimfeld K, Malanchini M, Krapohl E, *et al* (2018) The stability of educational achievement across school years is largely explained by genetic factors *NPJ Science of Learning* **3**: 16.

150 Plomin R (2018) *Blueprint: how DNA makes us who we are* (Cambridge, MA: The MIT Press).

151 Tanigawa Y, Qian J, Venkataraman G, *et al* (2022) Significant sparse polygenic risk scores across 813 traits in UK Biobank *PLoS Genetics* **18(3)**: e1010105.

152 Morris TT, Davies NM, and Davey Smith G (2020) Can education be personalised using pupils' genetic data? *Elife* **9**: e49962.

153 Murphy R and Wyness, G (2020) Minority report: the impact of predicted grades on university admissions of disadvantaged groups *Education Economics* **28(4)**: 333–50.

154 Benton T and Sutch T (2013) *Exploring the value of GCSE prediction matrices based upon attainment at Key Stage 2* (Cambridge: Cambridge Assessment).

So where does the science leave us? The performance of PGIs at the population level is expected to improve as the science advances and as PGIs are integrated with additional data. However, even with these developments, it seems likely that PGIs at best might serve as soft indicators to guide educational support and intervention, but not as deterministic predictors of individual educational outcomes. For instance, just as advances in medical **genomics** have allowed for early interventions in specific health conditions, PGIs might one day help educators and parents understand why a child is struggling with a particular skill, such as reading. Instead of a teacher simply noting that a child in Year 2 is taking longer than most to read well, PGI-informed insights might offer an additional piece of information that could help guide educational decision making and support.

Yet even a 'softer' approach to translation in educational settings brings with it a host of ethical considerations that cannot be overlooked. While identifying children more likely to experience educational difficulties – especially those related to special educational needs and difficulties (SEND) – is a valuable goal, potential future implementation of PGIs for this purpose could unintentionally worsen inequities. Given that SEND provision is limited and often depends on securing an Education, Health, and Care Plan (EHCP), for which a case must be made, if PGIs were ever used as evidence to support such cases, issues could arise about unequal access; with families who are more informed or resourced potentially better positioned to leverage this opportunity. Similar challenges are already evident for ADHD and dyslexia, where access to opportunities for diagnosis, and therefore support, may reflect socioeconomic disparities. Additional ethical concerns include the potential for discrimination (taking adverse action against a person or group based on their PGI) and stigmatisation (viewing a person or group as 'less' than others based on their PGI), as well as the emotional and behavioural impact of knowledge about one's PGI (the idea of self-fulfilling prophecies). Some of these concerns also apply to non-genetic measures used to predict future educational outcomes. For example, cognitive test scores at age 11 – far more accurate predictors of future achievement than PGIs – have been used in England to stream children into grammar schools. This practice remains highly controversial, as it relies on test scores measured with error to allocate children to entirely different educational systems. The same issues apply to PGIs, whether the intention is to allocate children to different educational settings or to identify those with specific educational needs. This highlights the broader challenges of ensuring that predictive measures in education, whatever their source, are implemented in ways that avoid reinforcing inequities or fostering deterministic views of children's potential.

Questions have also been raised about how ownership, privacy and anonymity would be safeguarded if PGIs were to be used for individual prediction, particularly in relation to complex data linkage streams that would be needed. For instance, the Information Commissioner's Office (ICO) – the UK's data protection and information rights body – has recently produced a **'Tech Futures' genomics report** that highlights challenges surrounding data security, transparency and consent, purpose of use, and implications of PGI information for other family members.

A particular ethical concern noted in this report and in the academic literature is the premature commercial application of PGIs for social and behavioural phenotypes via poorly regulated direct-to-consumer (DTC) companies (see **Section 5.3**).[155] As discussed in **Section 3.8**, this particular risk might be minimised by embedding unambiguous statements in research outputs about what cannot be done with PGIs, or by adding a Creative Commons 'Attribution Non-Commercial Share Alike 4.0 International (CC BY-NC-SA 4.0)' licence that would require permission to be sought from the lead researchers for use of the GWAS summary data.[156]

There are also worries that poor genomic literacy might lead to genetic determinism – the mistaken belief that a PGI alone determines an outcome, with little or no role for the environment or individual agency. Some studies have shown that while genomic literacy is improving over time in society, there is still much room for improvement.[157] Educational interventions to improve genomic literacy among the general public, school-aged children and educational professionals is widely recognised by the experts we spoke to and in the literature as an issue requiring urgent attention.[158]

Genomic information is set to become more mainstream in society in the coming years as genomic analyses become increasingly accessible to the public via DTC companies; as **Genomics England** begins testing babies for 200 rare treatable genetic conditions; and as **Our Future Health** communicates PGI results to research participants. We revisit some of these issues in **Section 5**. It is worth noting that many of these concerns are **also present in genomic healthcare research**.

4.2 PGIs as tools for basic research

While individual prediction using PGIs is a topic of much debate, a less contested application of educationally relevant PGIs is to incorporate them as research variables to enable genetically informed social science and educational research.

Such research requires genomic variation or PGI data for each research participant to be accessible to researchers. Unlike traditional genetically informed approaches, it does not require specialised study designs such as twin or adoption cohorts, although deeper insights can be gained by combining PGIs with family (sibling, parent–offspring) structures.[159] Additionally, because PGIs explain more variation than individual genetic variants, when a PGI is already available, statistically well-

155 De Hemptinne MC and Posthuma D (2023) Addressing the ethical and societal challenges posed by genome-wide association studies of behavioral and brain-related traits *Nature Neuroscience* **26(6)**: 932–41.

156 Ibid.

157 Little ID, Koehly LM, and Gunter C (2022) Understanding changes in genetic literacy over time and in genetic research participants *American Journal of Human Genetics* **109(12)**: 2141–51.

158 Asbury K, McBride T, and Rosie B (2022) Can genomic research make a useful contribution to social policy? *Royal Society Open Science* **9(11)**: 220873.

159 Harden KP and Koellinger PD (2020) Using genetics for social science *Nature Human Behaviour* **4**: 567–76.

powered studies can be performed with only a few hundred research participants. This feasibility of integrating PGIs into study designs is supported by the development of resources such as the **PGS Catalog**, the **Polygenic Index Repository** and the **PRS Knowledge Base**. These repositories allow researchers to calculate and evaluate PGIs in large research cohorts, and – when genomic data are available – in their own research datasets. The latter may be a source of issues, as generating PGIs from individual-level genetic data requires specialist skills and knowledge and is complicated by the variety of tools available.[160]

These methodological opportunities are being facilitated by the increasing availability of cohorts with both PGI and environmental measures that are, in some instances, securely linked to national education data. For example, in England's Millennium Cohort Study (MCS), PGIs (but not raw genomic data) can be securely imported into the **UK Data Service** SecureLab and linked to the Department for Education's **National Pupil Database (NPD)**. This allows MCS researchers to perform PGI-informed research that seeks to untangle the social, environmental and genomic contributions to differences in attainment, exclusion and pupil absences.

We highlight a selection of key applications from the literature in the following subsections, including research funded by the Nuffield Foundation. The utility of PGIs in social science research is a topic of some discussion, as evidenced by a recent review article[161] that garnered 24 academic commentaries reflecting a broad spectrum of views. One criticism is that PGIs can capture confounding (see **Section 2.8**) and may reflect effects mediated by or conditional on environmental factors. Consequently, PGIs do not offer a complete or 'clean' separation of genetic and environmental influences. Moreover, as discussed in **Section 4.1**, the predictive accuracy of PGIs varies, even within a genetically homogeneous population.

4.3 Partial control of genetic confounding in observational research

Identifying modifiable environmental factors that influence educational outcomes is a key objective for social scientists, as this knowledge can guide interventions. However, **gene–environment correlation** presents a challenge in this endeavour, as it suggests that associations between environmental factors and educational outcomes may be influenced by genetic factors (genetic **confounding**). This may complicate efforts to pinpoint environmental factors as causal points of intervention.

A range of established genetically informed methods can control for genetic confounding when examining links between environmental or social predictors and

160 Ni G, Zeng J, and Revez JA (2021) A comparison of ten polygenic score methods for psychiatric disorders applied across multiple cohorts *Biological Psychiatry* **90(9)**: 611–20.

161 Burt CH (2023) Challenging the utility of polygenic scores for social science: environmental confounding, downward causation, and unknown biology *Behavioral and Brain Sciences* **46**: e207.

child educational outcomes, such as twin, adoption, children-of-twin and sibling comparison studies.[162] PGIs offer a new approach that can partially account for, or at least allow researchers to acknowledge, the presence of genetic confounding in observational datasets. Specifically, researchers can assess whether the association between an environmental variable and an outcome weakens after accounting for participants' genetic propensities as indexed by a PGI for the outcome of interest.[163]

Pingault *et al* (2021) applied this approach in a UK twin cohort, using the EduYears PGI to examine the observed association between maternal education and child ADHD and educational achievement.[164] After taking the PGI into account, the strength of the relationships reduced. This finding does not negate the role of parental education in child development, but it suggests that a portion of the observed relationship (in this instance) is mediated by the genetic effects – and any other effects – captured by the EduYears PGI.

This has implications for social science research. Failing to account for genetic confounding might lead to misleading conclusions about the causal effects of environments,[165] wasting research resources and leading to ineffective interventions or even harm.[166] At a minimum, recognising and addressing genetic confounding can help researchers to more clearly and precisely identify the effects of environmental and social factors on child educational outcomes.[167] This approach is not perfect. PGIs only partially control for genetic confounding (due to **missing heritability**), and they capture environmental factors in ways that researchers are only starting to unpick. So, while controlling for a PGI will remove some genetic confounding, it may also remove some important environmental variation.

These complex issues underscore the growing importance of cross-disciplinary collaboration between behavioural and statistical geneticists, developmentalists and social scientists. For example, researchers with expertise in genomic data analysis and the development of PGIs may lack the in-depth knowledge of the educational system or epidemiological analyses that other scientists bring. Combining expertise across disciplines will help to ensure that basic genomic research is not only methodologically robust, but also relevant to real-world educational settings and child development.

---

162 Jami ES, Hammerschlag AR, Bartels M, *et al* (2021) Parental characteristics and offspring mental health and related outcomes: a systematic review of genetically informative literature *Translational Psychiatry* **11**: 197.

163 Pingault JB, Allegrini AG, Odigie T, *et al* (2022) Research review: how to interpret associations between polygenic scores, environmental risks, and phenotypes *Journal of Child Psychology and Psychiatry* **63(10)**: 1125–39.

164 Pingault JB, Rijsdijk F, Schoeler T, *et al* (2021) Genetic sensitivity analysis: adjusting for genetic confounding in epidemiological associations *PLOS Genetics* **17(6)**: e1009590.

165 Jaffee SR and Price TS (2012) The implications of genotype-environment correlation for establishing causal processes in psychopathology *Developmental Psychopathology* **24(4)**: 1253–64.

166 Hart SA, Little C, and van Bergen E (2021) Nurture might be nature: cautionary tales and proposed solutions *NPJ Science of Learning* **6(1)**: 2.

167 Harden KP and Koellinger PD (2020) Using genetics for social science *Nature Human Behaviour* **4**: 567–76.

4.4 Intergenerational transmission of educational risk and resilience

Just as the interdependence of genes and environments can bias estimates of environmental influences, it can also bias estimates from molecular genomic studies, such as GWASs. As discussed in **Section 2**, the common practice of using unrelated samples in a GWAS can induce links (correlations) between an individual's phenotype and their **genotype** that are non-causal. This can arise through several mechanisms, one of which is **gene–environment correlation**.

The increased availability of genomic datasets for families (such as siblings, or parent–child trios) is starting to provide new ways to examine which types of gene–environment correlation impact the detection and estimation of genetic effects on phenotypes that relate to education.[168] For instance, jointly modelling genomic and phenotypic data from families allows researchers to examine how phenotype-associated genetic variants that are not shared between parents and children still impact child outcomes via environmental routes.

It is apparent from these new approaches that the EduYears PGI contains a substantial portion of non-causal ('indirect') associations. The first study of this kind used genomic and phenotypic data from trios to demonstrate that genetic variants associated with years in education that were not inherited by the child still predicted the length of time that the child spent in education.[169]

This demonstrates that the parental EduYears PGI is capturing some aspects of the shared environment that correlate with child outcomes. In essence, the PGI captures not only genetic variants acting directly in the individual, but also indirect influences of parental genotypes on child phenotypes that are operating via environmental paths, referred to as '**genetic nurture**' or 'dynastic' effects.

A recent meta-analysis funded by the Nuffield Foundation found that about one-third of the variance accounted for by the EduYears PGI is due to genetic nurture effects.[170] Family-based methods are starting to be applied to a wider range of PGIs that relate to education in order to obtain estimates of genetic effects that are independent of genetic nurture and other confounding influences.[171]

168 Harden KP and Koellinger PD (2020) Using genetics for social science *Nature Human Behaviour* **4**: 567–76.

169 Kong A, Thorleifsson G, Frigge ML, *et al* (2018) The nature of nurture: effects of parental genotypes *Science* **359(6374)**: 424–8.

170 Wang B, Baldwin JR, Schoeler T, *et al* (2021) Robust genetic nurture effects on education: a systematic review and meta-analysis based on 38,654 families across 8 cohorts *American Journal of Human Genetics* **108(9)**: 1780–91.

171 McAdams TA, Cheesman R, and Ahmadzadeh YI (2023) Annual research review. Towards a deeper understanding of nature and nurture: combining family-based quasi-experimental methods with genomic data *Journal of Child Psychology and Psychiatry* **64**: 693–07.

An unresolved question is what the environmental 'genetic nurture' paths are. Incorporating parental phenotypes or measured environments into family-based PGI studies can help researchers to determine whether genetic nurture effects act through specific processes, such as parenting behaviours or the influence of extended family members.[172] For example, the Nuffield-funded Wang *et al* study reported that genetic nurture effects can largely be accounted for by parental education and socioeconomic status, providing new insights into how inequalities are transmitted within families.[173]

Another question is how much it matters that real-world indirect processes are being indexed by the EduYears PGI. If the goal is to understand the direct effects of an individual's own genetic make-up on education, accounting or controlling for indirect effects becomes important. For instance, it is possible that direct genetic effects might reveal a little more about the underlying biology of a phenotype.

One way to achieve this is by performing within-family GWASs that include at least two members of the same family.[174] However, this design requires genomic data from families (two generations, or siblings), which are harder to recruit than unrelated individuals. As a result, few large family-based GWASs of social and behavioural phenotypes that relate to education have been conducted. Those that have been conducted indicate that much of the prediction for EduYears and cognitive function PGIs comes from factors other than the direct effects of an individual's own genetic make-up.[175]

In contrast, if the goal is to maximise the predictive power of a PGI for a given population, it has been argued that biases and confounding from indirect effects do not necessarily need to be accounted for, as PGIs might still be useful even when biases are present.[176] For example, yellowed fingers are not a causal factor in lung disease but serve as a good predictor of smoking, which therefore provides information about a person's chances of developing lung disease. Similarly, some academics suggest that the additional predictive power contributed by indirect effects might be practically useful for identifying groups of individuals who are more likely to face challenges in education, where they have not already been identified as such.[177]

172 Nivard MG, Belsky DW, Harden KP, *et al* (2024) More than nature and nurture, indirect genetic effects on children's academic achievement are consequences of dynastic social processes *Nature Human Behaviour* **8(4)**: 771–8.

173 Wang B, Baldwin JR, Schoeler T, *et al* (2021) Robust genetic nurture effects on education: a systematic review and meta-analysis based on 38,654 families across 8 cohorts *American Journal of Human Genetics* **108(9)**: 1780–91.

174 Howe LJ, Nivard MG, Morris TT, *et al* (2022) Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects *Nature Genetics* **54**: 581–92.

175 Ibid; Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

176 Plomin R and von Stumm S (2022) Polygenic scores: prediction versus explanation *Molecular Psychiatry* **27**: 49–52.

177 Ibid.

4.5 Searching for gene–environment interactions

GWASs average out environmental influences in the samples studied, allowing genetic associations to be detected without directly considering environmental or educational contexts. However, given the known contribution of these factors to educational outcomes developmentally, it is important to understand whether genetic effects identified in a GWAS differ by, or are conditional on, environment: a phenomenon known as **gene–environment interaction** (GxE).[178]

GxE is conceptually distinct from gene–environment correlation. The latter is about the way that certain environments tend to occur with certain genotypes. GxE is concerned with how genes and environments combine multiplicatively so that if the environment is changed for a certain genotype (such as a low EduYears PGI), the outcome will be different. A classic example is the treatment of the **monogenic disorder** PKU, where altering the environment through a low-phenylalanine diet mitigates the genetic effect (see **Section 2.4**). In education, environments span various levels, including biological (e.g. epigenetic alterations, the microbiome), school (e.g. teaching quality or method, resources), family (e.g. parental income, stressful life events), neighbourhood (e.g. crime, access to green space, pollution), and broader societal factors.

GxE research is highly relevant to translation, as it offers insights into how modifiable environmental factors may differentially affect children with a genetically higher (or lower) likelihood of experiencing poorer educational outcomes.

Historically, environments have been easier to measure than genetic variation, limiting early studies to twin designs that assessed whether heritability estimates varied as a function of specific environments.[179] With the advent of molecular genomic studies (see **Section 2.5**), GxE research shifted towards looking for interactions between **candidate gene** variants and social and behavioural phenotypes. However, many candidate gene interaction studies failed to replicate due to a lack of **statistical power** (i.e. small sample sizes), and most reported candidate gene associations are now known to have been false.[180]

The advent of PGIs has reinvigorated GxE studies because – unlike single variants – PGIs can reliably predict a portion of the observed differences in phenotypes among children and adolescents.[181] To date, GxE research using PGIs that relate to education

178 Kendler KS and Eaves LJ (1986) Models for the joint effect of genotype and environment on liability to psychiatric illness *American Journal of Psychiatry* **143(3)**: 279–89.

179 Purcell S (2002) Variance components models for gene–environment interaction in twin analysis *Twin Research* **5(6)**: 554–71.

180 Dick DM, Agrawal A, Keller MC, *et al* (2015) Candidate gene–environment interaction research: reflections and recommendations *Perspectives on Psychological Science* **10(1)**: 37–59.

181 Wilding K, Wright M, and von Stumm S (2024) Using DNA to predict education: a meta-analytic review *Educational Psychology Review* **36**: 102.

is still in its early stages, and findings are mixed. For example, several studies,[182] including one funded by the Nuffield Foundation,[183] did not find interactions between PGIs and family environments in relation to child academic achievement. However, a recent, large Norwegian birth cohort study identified an interaction between students' EduYears PGI and school environment. This study found that the influence of students' PGIs on academic achievement differed by school, with higher-performing schools weakening the influence of a lower PGI.[184] Similarly, a US study found that students with a lower EduYears PGI were less likely to drop out of maths classes if they attended socioeconomically advantaged schools.[185]

The picture is similarly mixed for PGI-based GxE research into childhood mental health symptoms and neurodevelopmental conditions, with both significant[186] and non-significant interactions reported.[187]

These inconsistent findings have been attributed partly to the fact that PGI-based GxE studies require larger sample sizes than are typically available. For instance, in a Nuffield-funded study it was estimated that a minimum sample size of 4,000 would be needed to detect mid-sized interaction effects, rising to 75,000 for smaller effects.[188] Additionally, while many GxE studies use the same PGI, they examine different environmental factors measured in varying ways. This lack of consistency makes it difficult to directly compare findings across studies.[189]

Lastly, if the environments modelled are themselves more similar for genetically similar individuals (gene–environment correlation), the results can be misleading.[190] It is possible to control for such correlation in interaction studies by using family-based designs, but a particular challenge here, as previously noted, is the lack of large research cohorts that collect both genomic and phenotype data for related individuals.

---

182 Allegrini AG, Karhunen V, Coleman JRI, *et al* (2020) Multivariable G-E interplay in the prediction of educational achievement *PLoS Genetics* **16(11)**: e1009153.

183 Von Stumm S, Kandaswamy R, and Maxwell J (2023) Gene–environment interplay in early life cognitive development *Intelligence* **98**: 101748.

184 Cheesman R, Borgen NT, Lyngstad TH, *et al* (2022) A population-wide gene–environment interaction study on how genes, schools, and residential areas shape achievement *NPJ Science of Learning* **7**: 29.

185 Harden KP, Domingue BW, Belsky DW, *et al* (2020) Genetic associations with mathematics tracking and persistence in secondary school *NPJ Science of Learning* **5**: 1; a summary of Nuffield-funded GxE research using polygenic indices can be found here.

186 Nelemans SA, Boks M, Lin B, *et al* (2021) Polygenic risk for major depression interacts with parental criticism in predicting adolescent depressive symptom development *Journal of Youth and Adolescence* **50**: 159–76.

187 He Q and Li JJ (2021) A gene–environment interaction study of polygenic scores and maltreatment on childhood ADHD Research on *Child and Adolescent Psychopathology* **50**: 309–19.

188 Von Stumm S and Nancarrow AF (2024) New methods, persistent issues, and one solution: gene–environment interaction studies of childhood cognitive development *Intelligence* **105**: 101834.

189 Domingue BW, Trejo S, Armstrong-Carter E, and Tucker-Drob EM (2020) Interactions between polygenic scores and environments: methodological and conceptual challenges *Sociological Science* **7**: 465–86.

190 Keller MC (2014) Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution *Biological Psychiatry* **75(1)**: 18–24.

As genomic datasets continue to grow, these challenges will be addressed, although the timeline remains uncertain. Gene–environment interaction research was highlighted by some experts we spoke to as having the greatest potential for translation into education practice and policy, as it may identify which environments and interventions are most effective for different children.

4.6 Using polygenic indices to improve intervention studies

Many of the researchers we consulted noted issues with how educational intervention studies are designed and assessed. The primary point raised (and also noted in the literature) was that interventions do not work for every child, all of the time. By looking only for average effects of an intervention, children for whom the intervention is more or less effective – perhaps for genomic reasons – are not identified. To our knowledge, PGIs have not yet been incorporated into educational trials from the outset.

Another point raised is that as some PGIs, such as the EduYears PGI, predict a reasonable portion of phenotype variability, they can be used as control variables in educational **randomised controlled trials** and intervention studies to account for uninformative 'noise' in the data.[191] By reducing heterogeneity in this way – which is otherwise difficult to capture – statistical power is boosted, allowing such studies to be conducted using fewer research participants. A rough calculation by Meyer *et al* (2023) on two US intervention studies estimated that including an EduYears PGI as a control variable could allow for sample sizes to be reduced by 6.7%, even after accounting for the costs of collecting genomic data. [192]

The authors note that expected resource savings will depend on how much additional variance the PGI accounts for, which will differ depending on the specific PGI used. Additionally, the poor transferability of PGIs across genetic ancestries will present challenges. It seems likely that collecting genomic data as an inclusion criterion in educational trials could hinder participant recruitment and raise ethical concerns, such as ones around data privacy, that would need to be addressed.[193]

191 Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–49.

192 bid.

193 Fahed AC, Philippakis AA, and Khera AV (2022) The potential of polygenic scores to improve cost and efficiency of clinical trials *Nature Communications* **13**: 2922.

4.7 Mendelian randomisation: using genetic information to query causal relationships

Research based on observational data – meaning studies that analyse information collected without manipulation of conditions or conducting controlled experiments – has demonstrated that educational measures correlate with many important health and economic outcomes.[194] As discussed, disentangling cause and effect from observational data is difficult (see **Section 2.8**). For example, consider the replicated correlation between adolescent depression and later educational outcomes, where at a population level, depression can precede failure to complete compulsory schooling and lower grades.[195] Other unknown or unmeasured factors, such as socioeconomic status, may causally impact both school outcomes and likelihood of depression, making it look like there is a relationship between the two when there is not (i.e. socioeconomic status is a confounder). Alternatively, educational difficulties may make it more likely that an individual experiences depression symptoms (i.e. reverse causation), although this can be examined using longitudinal data. Evidence of the direction of causality between a risk factor (also referred to as an 'exposure') and a phenotype of interest (outcome) is a requirement for effective intervention strategies and estimation of their impact.[196]

Genomic data can potentially be used to strengthen observational data and draw more reliable conclusions about cause and effect. One approach is **Mendelian randomisation** (MR), which parallels the principles of randomised controlled trials. Instead of random allocation to conditions, MR uses genetic variants associated with the risk factor of interest to avoid confounding and reverse causation (see **Figure 4**).

MR works by taking advantage of the fact that **alleles** of a genetic variant are randomly inherited from parent to child. Because inheritance of DNA variation is unrelated to environmental factors, these factors by definition cannot be confounders, unlike in traditional observational research. Furthermore, because alleles are fixed at conception, the possibility of reverse causation is eliminated.[197]

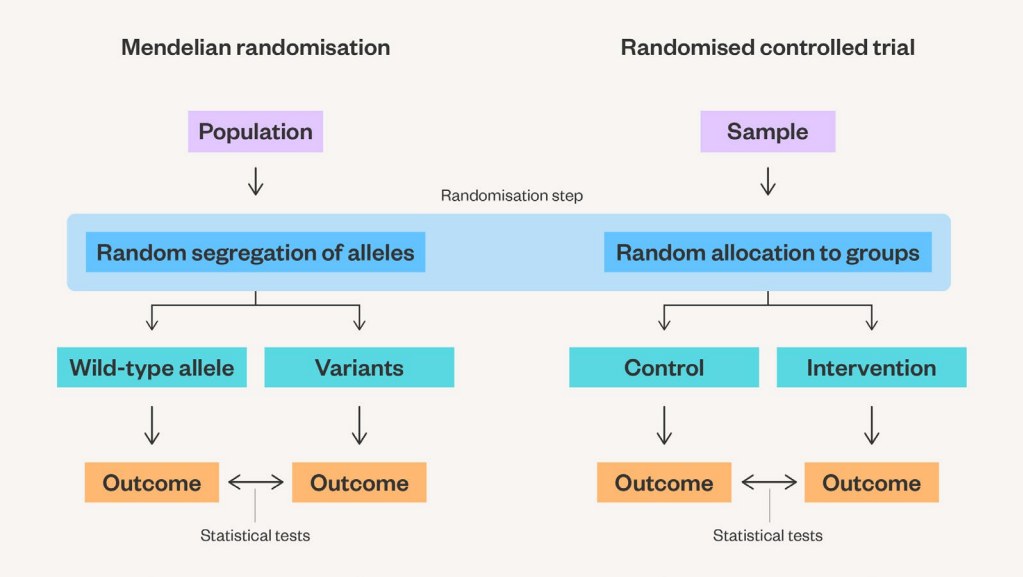
194 Farquharson C, McNally S, and Tahir I (2024) Education inequalities *Oxford Open Economics* **3(1)**: i760–820.

195 Riglin L, Petrides KV, Frederickson N, and Rice F (2014) The relationship between emotional problems and subsequent school attainment: a meta-analysis *Journal of Adolescence* **37**: 335-46.

196 Pingault JB, O'Reilly PF, Schoeler T, *et al* (2018) Using genetic data to strengthen causal inference in observational research *Nature Reviews Genetics* **19**: 566–80.

197 Chen LG, Tubbs JD, Liu Z, Thach TQ, and Sham PC (2024) Mendelian randomization: causal inference leveraging genetic data *Psychological Medicine* **54(8)**: 1461–74.

Figure 4: An overview of Mendelian randomisation



In the MR method, genetic variants must be associated with the risk factor, but not with any other factors that affect the outcome. This ensures that any association between the genetic variant and the outcome must occur only through the variant's relationship with the risk factor, implying a causal effect of the risk factor on the outcome.[198]

For example, suppose a simplified hypothetical scenario where individuals with more copies of the A allele for a variant associated with years spent in education (but are not associated with depression) stay in education for longer than individuals with fewer copies of the A allele (i.e. 2 > 1 > 0). If individuals with more copies of the A allele report fewer depression symptoms, then researchers can estimate the causal effects of years in education on depression. The effect of the education-associated variant on depression is operating only through education, implying that more time spent in education causally decreases risk for depression.

MR has been successfully applied in medical contexts to demonstrate causal relationships in health, for example between blood lipid levels and myocardial infarction.[199] It is also starting to be used to query the causal relationship between years spent in education and later life outcomes such as Alzheimer's disease.[200]

198 Burgess S, Davey Smith G, Davies NM, *et al* (2023) Guidelines for performing Mendelian randomization investigations: update for summer 2023 Wellcome *Open Research* **4**: 186.

199 Voight BF, Peloso GM, and Orho-Melander M (2012) Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study *Lancet* **380(9841)**: 572–80.

200 Seyedsalehi A, Warrier V, Bethlehem RAI, Perry BI, Burgess S, and Murray GK (2023) Educational attainment, structural brain reserve and Alzheimer's disease: a Mendelian randomization analysis *Brain* **146(5)**: 2059–74.

However, in order for valid conclusions to be drawn from a Mendelian randomisation analysis, several assumptions must be met. Specifically, MR requires that (1) the genetic variant robustly associates with the risk factor, (2) the genetic variant is not associated with the outcome via a confounding path, and (3) the variant does not affect the outcome directly, only indirectly via the risk factor[201] (see Figure 5).

This last assumption can be difficult to meet because (as described in **Section 3.2**) many of the variants identified in a GWAS are **pleiotropic** and associate with more than one phenotype. If genetic **pleiotropy** is 'horizontal' (the variant associates with the outcome via paths independent of the risk factor), the assumption is violated.

Figure 5: The assumptions for Mendelian randomisation

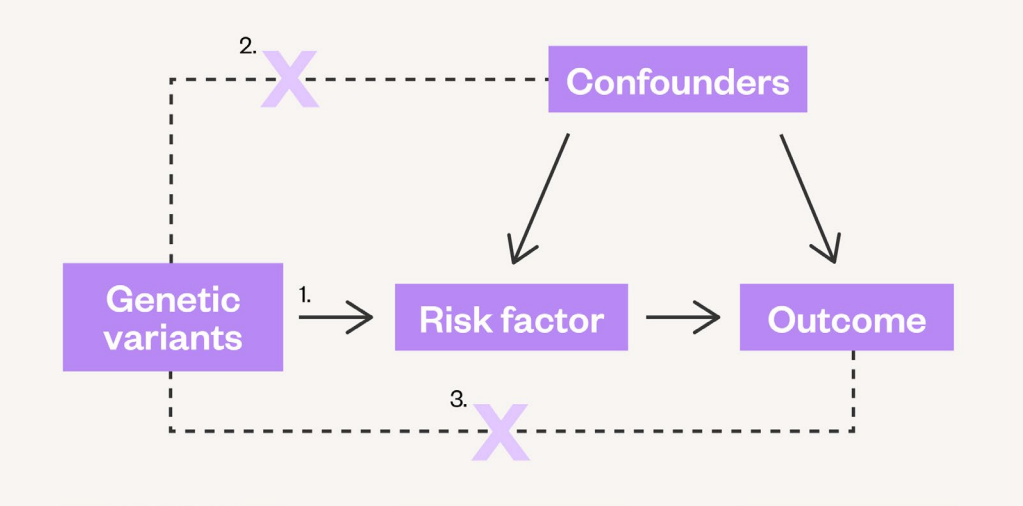


Figure 5 illustrates that Mendelian randomisation aims to estimate the causal effect of one variable (the risk factor) on another (the outcome) using a genetic variant ('genetic instrument') to control for any unmeasured confounders.

If the assumptions of MR are not met, results may be biased,[202] although newer developments in MR methods can relax or test these assumptions.[203] Further, it is unclear how the presence of **assortative mating** and genetic nurture effects complicate the interpretation of MR results.[204] Both effects are present for genetic variants associated with years spent in education.

201 This is vertical pleiotropy, where the association of the variant with a phenotype is completely mediated by an intermediary phenotype. Horizontal pleiotropy is where the variant operates through multiple pathways, which violates Mendelian randomisation assumptions, making interpretation of results difficult.

202 De Leeuw C, Savage J, Bucur IG, Heskes T, and Posthuma D (2022) Understanding the assumptions underlying Mendelian randomization *European Journal of Human Genetics* **30(6)**: 653–60.

203 Burgess S and Thompson SG (2017) Interpreting findings from Mendelian randomization using the MR-Egger method *European Journal of Epidemiology* **32(5)**: 377–89.

204 Brumpton B, Sanderson E, Heilbron K, *et al* (2020) Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses *Nature Communications* **11**: 3519.

Some of the researchers we spoke to think that the addition of MR analyses to social scientists' analytical toolkits will usefully contribute to educational epidemiological research. However, researchers also cautioned that this method requires extremely careful study design and interpretation, because it's often unclear whether all assumptions can be met – especially for complex, pleiotropic traits like years in education or educational achievement

5 Implications for research, education practice and policy

In the previous section we described how PGIs are being applied in basic research, and the scientific and ethical limitations of harnessing PGIs for individual prediction. In researching this report, we also uncovered several data gaps that have not yet been described.

In this section, we consolidate these insights to highlight key open questions and limitations, why they matter for translation, and some of the ethical implications of leaving them unresolved. We also hope to stimulate thinking about research priorities and the identification of plausible, ethical routes from basic science to meaningful translation, while highlighting both the risks and opportunities that lie ahead.

5.1 Polygenic indices as research tools for translation

As outlined in **Section 4.2**, the application of PGIs as research tools presents opportunities to unpick gene–environment correlations, deepen understanding of environmental causes, and explore the scope for changing educational outcomes. In the medium-to long-term, these approaches may inform social and educational policies and practice. However, the use of PGIs in a research context is not without challenges: scientific, practical and ethical.

In addition to the diversity problem, PGIs do not capture the entire genomic influence on a **phenotype**, making them an imperfect tool to control for **genomic confounding**. Additionally, PGIs may index not only the direct effects of an individual's own **genotype** but the indirect effects of their genetic relatives and demography. While not a perfect solution, family genomic data (siblings or parent–child trios) can help to distinguish between direct and indirect genetic effects and better control for other sources of **confounding**, such as **population stratification**.[205] However, large family-based genomic resources remain scarce.

205 Howe LJ, Nivard MG, Morris TT, *et al* (2022) Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects *Nature Genetics* **54**: 581–92.

The need for large sample sizes to detect associations for highly polygenic phenotypes often leads to researchers relying on poorly measured phenotypes, as they must default to the lowest common denominator of phenotype when harmonising across cohorts. A related issue raised by some researchers is that GWAS efforts tend to prioritise ever larger studies of phenotypes above improving the ways they are measured. Further, the purpose of education is much broader than the accumulation of knowledge and certificates: phenotypes such as resilience, communication skills, self-regulation and curiosity,[206], as well as behavioural problems and conditions such as motor disorders are also important – but comparatively underexamined.[207]

Researchers are tackling these data gaps by employing novel statistical methods using existing datasets, but these approaches are technically complex to perform and interpret.[208] Ultimately, once GWAS sample sizes have been maximised using existing data, better and more diverse phenotype measurements will be needed for deeper and more holistic insights.

One particular challenge lies in collecting high-quality, detailed phenotypic data in sample sizes large enough to identify phenotype-associated genetic variants.[209] This is being addressed through the development of gamified online cognitive tests,[210] phenotype imputation,[211] the use of data collected from personal digital devices ('digital phenotyping') and sensing technologies[212], and globally coordinated efforts to gather detailed phenotype data (e.g. see **GenLang**). However, careful consideration of the implications for data privacy and surveillance is needed.[213]

Until this data gap is bridged, there is a risk that genetically informed research will overlook vital aspects of behaviour and functioning that influence education and later life outcomes, such as employment. This may perpetuate inequalities by prioritising narrow measures of attainment and cognitive skills over 'softer' transferrable skills, such as teamwork and interpersonal communication, that are of value to employers and society.

206 Kautz T, Heckman JJ, Diris R, Ter Weel B, and Borghans L (2014) Fostering and measuring skills: improving cognitive and non-cognitive skills to promote lifetime success, OECD Education Working Paper 110 (Paris: OECD Publishing).

207 Gidziela A, Ahmadzadeh YI, Michelini G, *et al* (2023) A meta-analysis of genetic effects associated with neurodevelopmental disorders and co-occurring conditions *Nature Human Behaviour* **7(4)**: 642–56.

208 Demange PA, Malanchini M, Mallard TT, *et al* (2021) Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction *Nature Genetics* **53**: 35–44.

209 Yeatman JD, Tang KA, Donnelly PM, *et al* (2021) Rapid online assessment of reading ability *Scientific Reports* **11**: 6396.

210 Malanchini M, Rimfeld K, Gidziela A, *et al* (2021) Pathfinder: a gamified measure to integrate general cognitive ability into the biological, medical, and behavioural sciences *Molecular Psychiatry* **26**: 7823–37.

211 Dahl A, Thompson M, An U, *et al* (2024) Phenotype integration improves power and preserves specificity in biobank-based genetic studies of major depressive disorder *Nature Genetics* **55(12)**: 2082–93.

212 Elbaum B, Perry LK, and Messinger DS (2024) Investigating children's interactions in preschool classrooms: an overview of research using automated sensing technologies *Early Childhood Research Quarterly* **66**: 147–56.

213 Perez-Pozuelo I, Spathis D, Gifford-Moore J, Morley J, and Cowls J (2021) Digital phenotyping and sensitive health data: implications for data governance *Journal of the American Medical Informatics Association* **28(9)**: 2002–8.

A related challenge is that environmental measures available in large genomic cohorts and biobank resources are often inadequate, typically relying on demographic data and self-reported aspects of the home environment.[214] This lack of precision and breadth partly reflects the methodological and resource challenges of capturing the full range of social and environmental influences throughout development, in sufficient detail and at scale.[215] Experts we consulted noted inconsistent measurement of school environments, especially in relation to what happens in classrooms, the support and learning resources available across schools, and the type of instruction provided. It is unclear how this school-level data gap might be solved without linking to existing high-quality data or the collection of new prospective studies.

Ultimately, a more integrated evaluation of home, school, and wider environments and experiences across the lifespan will require collaboration across research disciplines, particularly when incorporating PGIs that can be challenging to causally interpret. This underscores the importance of multidisciplinary collaboration that draws on expertise from ethicists and the social, epidemiological, genomic and developmental sciences.[216]

5.2 PGIs as a tool for individual-level prediction

There is ongoing debate in the scientific community about the translational potential of PGIs in educational practice and policy.[217] Some argue that genomic prediction could help to guide classroom practices; to better meet students' specific needs and improve their educational experiences and outcomes.[218] Typically, this could involve identifying students who are more vulnerable to falling behind academically, enabling early targeted intervention to support learning.

However, while some PGIs relevant to education are powerful for group-level predictions and relationships, they are poor for individual-level predictions. The PGIs of high academic achievers are, on average, higher than those of other students, but the PGI distributions of these two groups heavily overlap. PGIs in isolation are thus poor at predicting where a specific student will fall.

One concern is distinguishing between 'identifying students at risk' and 'prediction'. Prediction implies a degree of certainty about a student's likelihood of a particular

214 Von Stumm S, Kandaswamy R, and Maxwell J (2023) Gene–environment interplay in early life cognitive development, *Intelligence* **98**: 101748.

215 Von Stumm S and d'Apice K (2022) From genome-wide to environment-wide: capturing the environome *Perspectives in Psychological Science* **17(1)**: 30–40.

216 Pingault JB, Allegrini AG, Odigie T, *et al* (2022) Research review: how to interpret associations between polygenic scores, environmental risks, and phenotypes *Journal of Child Psychology and Psychiatry* **63(10)**: 1125–39.

217 Meyer MN, Appelbaum PS, Benjamin DJ, *et al* (2023) Wrestling with social and behavioral genomics: risks, potential benefits, and ethical responsibility *Hastings Centre Report* **53(1)**: S2–4.

218 Plomin R and von Stumm S (2018) The new genetics of intelligence Nature Reviews Genetics **19(3)**: 148–59; Plomin R (2018) *Blueprint: how DNA makes us who we are* (Cambridge, MA: The MIT Press).

outcome, whereas identifying risk suggests more that further investigation or support may be required. These are similar but distinct concepts, and in practice there is a risk of conflation. This could result in PGIs being misused to make assumptions about a student's potential or to categorise students incorrectly. While this does not mean that PGIs should be dismissed entirely, it highlights a real conundrum that will need to be carefully navigated in real-world situations.

It is useful to consider how other indicators of student achievement have been used in schools, often with unintended consequences. For example, contextualised value-added (CVA) models for assessing school performance were discontinued due to concerns that they embedded lower expectations for students with specific contextual variables, such as eligibility for free school meals.[219] This led to a shift towards using prior attainment as the basis for targeting support, rather than focusing on background characteristics that predict low achievement on average. Similarly, it is possible to envisage scenarios where PGIs might be misused in real-world settings to limit achievement expectations, or to incentivise schools to deprioritise students with lower PGIs.

As already raised, PGIs do not predict equally well across diverse populations, posing the risk of entrenching or exacerbating educational and social inequalities. Moreover, the EduYears PGI captures environmental and social influences raising ethical concerns about making decisions based on not only a student's genetic make-up, but also other factors such as social advantage. Within-family GWAS-derived PGIs, which control for some of these confounders, may reduce this concern, but the PGI will explain less variance, making it less accurate for individual-level prediction.[220]

A key question is whether PGIs offer meaningful benefits beyond non-genomic predictors, such as prior achievement or familial factors, aside from the fact that they are, in principle, available prior to the start of schooling and are specific to siblings within the same family. While the predictive power of PGIs may improve in the near future, the PGI for years in education is already approaching its theoretical maximum based on heritability estimates.[221] Additionally, due to **pleiotropy**, a PGI for one phenotype may predict a second. As a result, PGIs lack specificity, and calculating a student's PGI for a particular phenotype can also reveal information about other genetic vulnerabilities or strengths. It may also enable secondary inferences to be made about family members, such as relatedness or other sensitive characteristics.

If the goal of individual prediction is to inform intervention, PGIs face another limitation: they do not map neatly to specific biological processes or pathways, nor do they provide insights into causal mechanisms. This means that it may not always

219 Dearden L and Vignoles A (2011) Schools, markets and league tables *Fiscal Studies* **32**: 179–86.

220 Okbay A, Wu Y, Wang N, *et al* (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals *Nature Genetics* **54**: 437–49.

221 Ibid.

be immediately clear how and when to act in order to mitigate potential difficulties effectively. Related to this, the education sector may lack the tools and resources needed to support PGI-based intervention or support, such as the personalisation of educational provision.

Additional ethical concerns include the risk of embedding systemic – or creating new forms of – discrimination and stigmatisation, including perhaps for phenotypes that are not currently considered protected characteristics under UK equalities legislation. Inaccurate inferences based on PGIs may also be difficult for individuals to contest due to their technical complexity, and the probabilistic nature of prediction can be difficult to communicate effectively.[222] Further, receiving PGI information could negatively affect students' wellbeing and parents' and teachers' perceptions and expectations of students, with evidence showing that teacher perceptions can influence students' educational trajectories.[223] Concerns were also raised by the experts we consulted that the introduction of PGIs into practice and policy may detract from other efforts to improve education and address social and economic inequities.

5.3 Wider societal backdrop

These debates are unfolding in a context where much of the data needed to construct PGIs for educationally relevant phenotypes are publicly available, and the science continues at pace. As a result, there is growing awareness of the potential ethical and policy implications of PGIs being applied outside of research settings. For instance, corporations, such as those selling insurance and financial services, may see economic value in PGIs,[224] and a growing body of literature has highlighted the ethical implications of marketing and selling PGI reports – predicting both medical and non-medical phenotypes – directly to adults,[225] or to *in vitro* fertilisation (IVF) patients to inform embryo transfer decisions (known as PGT-P).[226]

222 Wallingford CK, Kovilpillai H, Jacobs C, *et al* (2023) Models of communication for polygenic scores and associated psychosocial and behavioral effects on recipients: a systematic review *Genetic Medicine* **25(1)**: 1–11.

223 Koivuhovi S, Jung A, Kilpi-Jakonen E, Little TD, and Vainikainen MP (2025) Influence of track placement and teachers' perceptions of children's academic schoolwork skills on the development of children's motivational self-beliefs and achievement *Teaching and Teacher Education* **153**: 104847.

224 Meyer MN, Papageorge NW, Parens E, *et al* (2024) Potential corporate uses of polygenic indexes: starting a conversation about the associated ethics and policy issues *American Journal of Human Genetics* **111(5)**: 833–40.

225 Onstwedder SM, Jansen ME, Cornel MC, and Rigter T (2024) Policy guidance for direct-to-consumer genetic testing services: framework development study *Journal of Medical Internet Research* **26**: e47389; Martins MF, Murry LT, Telford L, *et al* (2022) Direct-to-consumer genetic testing: an updated systematic review of healthcare professionals' knowledge and views, and ethical and legal concerns *European Journal of Human Genetics* **30**: 1331–43.

226 Grebe TA, Khushf G, Greally JM, *et al* (2024) Clinical utility of polygenic risk scores for embryo selection: a points to consider statement of the American College of Medical Genetics and Genomics (ACMG) *Genetics in Medicine* 26(4): 101052.

PGT-P is illegal in the UK under the Human Fertilisation and Embryology Act 2008 but is permissible in the USA. Several private US companies offer PGT-P as a service, including for intelligence.[227] Despite evidence that using PGT-P to screen IVF embryos for years in education or intelligence would have a limited impact,[228] in the US there is a degree of interest in, and moral acceptance of, doing so.[229]

In the UK, PGI reports are available to individuals outside of healthcare or educational settings via direct-to-consumer (DTC) companies.[230] While we are not aware of UK-based DTCs offering PGI reports for psychiatric disorders, years in education or intelligence, European companies already market PGI reports for social and behavioural phenotypes (e.g. **Gene Plaza**) and US company **Nucleus** recently launched genetic tests for intelligence and psychiatric conditions.

The growing DTC market makes it plausible that practical considerations such as low cost and accessibility may drive societal adoption of PGIs, irrespective of academic, scientific and ethical concerns. Will parents 'vote with their wallets' by obtaining PGI reports for their children? A PGI report indicating a higher likelihood of social, emotional or behavioural difficulties might prompt parents to approach their child's school to request assistance and support[231], or to appeal local authority SEND decisions. In such scenarios, how should the state respond, and what criteria should guide its actions? Schools are already **struggling to support** children with specific learning needs. Without additional funding, they may lack the capacity to provide support even if evidence existed for effective interventions.

5.4 Next steps: exploring the ethical unknowns

This report has touched on some of the ethical issues arising in both educational genomics research itself and its potential translation. However, much remains underexplored and requires further deliberation on the range of ethical challenges arising, and their potential impact on research, practice and policy contexts.

A number of the key ethical challenges arising from the research itself – such as those around insufficient diversity in datasets – apply across health and educational genomics research alike, and have been the subject of attention in academic literature. Alongside greater interdisciplinarity in ethical discourse, and consideration of the

227 Turley P, Meyer MN, Wang N, *et al* (2021) Problems with using polygenic scores to select embryos *New England Journal of Medicine* **385(1)**: 78–86.

228 Karavani E, Zuk O, Zeevi D, *et al* (2019) Screening human embryos for polygenic traits has limited utility *Cell* **179(6)**: 1424–35.

229 Meyer MN, Tan T, Benjamin DJ, Laibson D, and Turley P (2023) Public views on polygenic screening of embryos *Science* **379(6632)**: 541–3.

230 Park JK and Lu CY (2023) Polygenic scores in the direct-to-consumer setting: challenges and opportunities for a new era in consumer genetic testing *Journal of Personal Medicine* **13(4)**: 573.

231 Martschenko D, Trejo S, and Domingue BW (2019) Genetics and education: recent developments in the context of an ugly history and an uncertain future *AERA Open* **5(1)**.

ethical issues raised in addressing knowledge gaps and methodological limitations, further focus might be given to identifying and exploring current and future challenges unique to educational genomics research; what their impact might be on both research integrity and translation potential; and how they might be effectively addressed.

Issues arising from the potential translation of research findings into policy and practice have not yet been the subject of detailed ethical debate, and this represents a significant gap in extant knowledge. There has been little exploration of the ethical issues arising from translation of PGIs at a classroom level and, accordingly, whether they should be translated into education at all. A greater understanding of these areas and their implications for policy would likely be of benefit not only to education professionals and policymakers, but also to the scientific community in developing the scope of future genomic research.

Annex 1 The structure of DNA

A **DNA** molecule is made up of sugar residues, phosphate groups and bases. A base-sugar-phosphate group is called a nucleotide. DNA is comprised of two chains, each made of nucleotides. The two chains form a double helix held together by weak hydrogen bonds between opposed bases, with 'A' always pairing with 'T', and 'C' pairing with 'G'. This is called the 'Watson-Crick' base pairing rule. See Figure 6. The Watson-Crick base pairing is key to the DNA copying mechanism needed to duplicate DNA prior to cell division (mitosis).

To make a distinction between the two strands of DNA, one strand is called the 'forward' strand and the other the reverse strand. This means that when talking about DNA sequence – the order of As, Gs, Cs and Ts – only one strand is referred to, typically the forward strand.

Figure 6: DNA double helix.

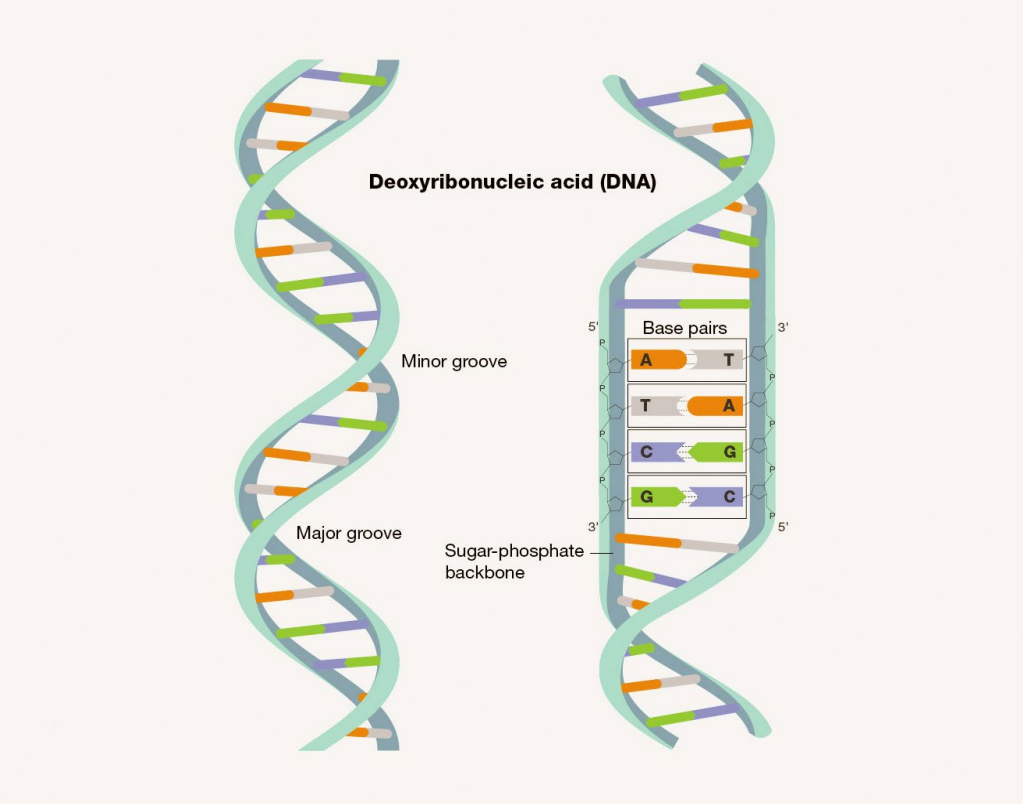


Image courtesy of the National Human Genome Research Institute Home | NHGRI

The human body is made up of gamete cells (sperm or egg cells) and somatic cells (all other cells of the body). Each somatic cell is diploid, which means it contains two copies of the **genome** – one inherited from each parent. The human genome sequence is approximately 3.1 billion nucleotides in length and organised into 23 pairs of **chromosomes**. These include 22 pairs of autosomes, numbered from largest (chromosome 1) to smallest (chromosome 22) based on their size, and one pair of sex chromosomes (XX for females and XY for males). One chromosome in each pair is inherited from each parent.

Gamete cells are haploid, which means they contain only one copy of the genome; 22 homologous chromosomes and a Y or a X chromosome. Gamete cells are formed by meiosis. During meiosis, the paired maternal and paternal chromosomes are shuffled (a process called recombination) resulting in a gamete that is a unique mixture of the parental genomes.

Annex 2 Genetic variation and genotyping methods

Genetic variation – changes in the sequence of DNA – can arise via copying errors during cell division, exposure to environmental mutagens (such as ionising radiation) and viral infection. This variation can occur in gamete (sperm and egg) cells or somatic (all other) cells, but only variation that arises in gametes can be passed on to offspring (i.e. inherited).

Evolutionary and demographic factors such as genetic drift, selection, migration and population size will influence the frequency of genetic variants in a population. Generally, genetic variants that are common in a population tend to be older (i.e. mutations that arose many generations ago) and are often neutral or beneficial to an individual's health and functioning. In contrast, rare variants are more likely to have deleterious effects on health, and tend to have arisen more recently, and may even be specific to a family or an individual.

Genetic variation ranges in scale. It can encompass large structural alterations, such as changes in the number or arrangement of chromosomes, or the number of copies of a segment of DNA (**called copy number variation, or CNVs**). The most frequent form of genetic variation in the human genome is a single base change where one base (A, G, C or T) is replaced by another. These are known as **single nucleotide polymorphisms, or SNPs** (pronounced 'snips') for short.

The term **allele** is used to refer to the multiple versions of a DNA sequence that exist for a genetic variant. In the case of SNPs, it simply refers to the different bases (A, G, C, or T) that can occur at that specific site. The combination of alleles for a genetic variant is called a **genotype**. For example, at a SNP where the alleles are C or T, an individual might have the genotype TC, TT, or CC. See **Figure 7**.

To date, genomic research on complex phenotypes that relate to education has largely focused on common genetic variation, such as SNPs. This is because they are easier than rare variants to detect and collectively account for a big proportion of genetic differences between people; a typical genome differs from the reference human genome sequence at roughly 3.5 million SNPs.[232] Because each SNP usually has only two possible **alleles**, upwards of 600,000 SNPs can be genotyped in large numbers of individuals using **SNP microarray** technology. This technology is

232 1000 Genomes Project Consortium (2015) A global reference for human genetic variation *Nature* **1;526(7571)**:68-74.

cheaper and faster than **DNA sequencing**-based approaches to genotyping for the whole genome (**whole genome sequencing**; WGS) or the protein coding regions of the genome (**whole exome sequencing**; WES).

Due to their low cost and speed, SNP microarrays are the primary technology used to perform **genome-wide association studies** (GWAS).

Figure 7: A single nucleotide polymorphism. Note the DNA sequence from one strand of the DNA helix only is shown, for both the maternally and paternally inherited genomes.

Annex 3 Heritability and common misconceptions

Heritability is a numerical estimate of the degree to which variation in a **phenotype** (such as height or reading ability) in a population is due to genetic variation between individuals in that population.

Historically, heritability of a phenotype was estimated using **twin studies** that exploit the fact that monozygotic (identical) and dizygotic (non-identical) twins share the same *in utero* and familial experiences, but differ in genetic similarity. Monozygotic twins develop from a single fertilised egg splitting into two embryos and therefore share the same genetic information, while dizygotic twins develop from two separately fertilised eggs and are no more genetically alike than full siblings. If pre- and post-natal environments are similar for both monozygotic and dizygotic twins (the equal environments assumption),[233] heritability can be estimated by the extent to which monozygotic twins are more similar for a phenotype than dizygotic twins by virtue of their higher genetic similarity.

More recently, analytical methods have been developed to estimate heritability using genotype data from thousands of individuals.[234] These methods – including Linkage Disequilibrium Score Regression (LDSC)[235] and Genome Relatedness Restricted Maximum Likelihood (GREML)[236] – measure the degree to which the genotyped common genetic variants that are spread throughout the genome contribute to a phenotype. Estimates from these approaches are called SNP-based heritability. LDSC and GCTA methods can also be used to estimate **genetic correlations** across phenotypes using genotype or GWAS summary data.

Both twin and SNP-based studies have shown that years spent in education and educational achievement are heritable. Genomic variation also contributes to many other phenotypes that impact education, such as general cognitive function

233 Richardson K and Norgate S (2005) The equal environments assumption of classical twin studies may not hold *British Journal of Educational Psychology* **75**:339–50.

234 Barry CJS, Walker VM, Cheesman R, *et al* (2023) How to estimate heritability: a guide for genetic epidemiologists *International Journal of Epidemiology* **52(2)**:624–32.

235 Bulik-Sullivan B, Loh PR, Finucane H, *et al* (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies *Nature Genetics* **47**:291–95.

236 Yang J, Benyamin B, McEvoy BP, *et al* (2010) Common SNPs explain a large proportion of the heritability for human height *Nature Genetics* **42(7)**:565-9.

(intelligence), psychiatric conditions such as depression and anxiety, neurodevelopmental conditions such as autism and ADHD, and specific learning difficulties such as dyslexia.[237]

Heritability is a statistic that is often misunderstood.[238] Here are some common misinterpretations and clarifications:

Misconception 1: Heritability estimates for a phenotype are static

The relative contribution of genetic and environmental factors to a phenotype are context-specific. As a result, twin and SNP-based heritability estimates can fluctuate depending on population-specific cultural, social and demographic factors, as well as how the phenotype is defined and measured, and the age at which it is assessed.

For example, the heritability of educational attainment differs across birth cohorts and countries[239] and can even change within the same population over time as environmental factors – such as educational policies and social context – change.[240] Generally, heritability of educational phenotypes tends to increase when equality of learning opportunities and environments improve. For instance, if every child had equally optimal education-relevant environments, genetic differences would become the primary source of variability in educational outcomes, and so heritability estimates would be very high. In contrast, in a country with limited educational resources or widespread poverty, environmental factors would be the limiting factor and heritability estimates would be low. Similarly, in countries with highly variable schools and home environments, these variable conditions would mainly account for differences in outcomes and result in low heritability estimates.

Heritability can also change across the lifespan. For instance, the twin-based heritability of general cognitive ability increases from ~40% in childhood to 66% in young adulthood.[241] This is thought to be due to gene–environment correlation, whereby individuals seek out and create environments that align with their genetic propensities as they grow up.

237 Polderman T, Benyamin B, de Leeuw C, *et al* (2015) Meta-analysis of the heritability of human phenotypes based on fifty years of twin studies *Nature Genetics* **47**:702–09.

238 Visscher P, Hill W, Wray N, *et al* (2008) Heritability in the genomics era – concepts and misconceptions *Nature Reviews Genetics* **9**:255–66.

239 Silventoine K, Jelenkovic A, Sund R, *et al* (2020) Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts *Scientific Reports* **10**:12681.

240 Rimfeld K, Krapohl E, Trzaskowski M, *et al* (2018) Genetic influence on social outcomes during and after the Soviet era in Estonia *Nature Human Behaviour* **2(4)**:269-75.

241 Haworth CM, Wright MJ, Luciano M, *et al* (2010) The heritability of general cognitive ability increases linearly from childhood to young adulthood *Molecular Psychiatry* **15(11)**:1112-20.

Misconception 2: Heritability tells you something about the sources of between-group differences

High heritability does not mean that differences in the average phenotype between two populations are due to genetic factors. Heritability measures variation within a group and provides no information about the size or sources of differences in phenotype between groups.

Misconception 3: Heritability provides information on the size of genomic effect for an individual

Heritability can sometimes be confused with genetic inheritance. Genetic inheritance is the passing of genes from parent to child. Heritability is what makes people in a population phenotypically different from each other – it does not explain what proportion of an individual's phenotype is inherited. For example, across the lifespan, the heritability of intelligence is ~ 50%.[6] That is, genetic differences can account for, on average, 50% of the differences in intelligence that can be measured between individuals. It does not mean that inherited genetic variation accounts for 50% of a given person's intelligence.

Misconception 4: Finding high heritability for a phenotype means it is immutable

High heritability means that a high proportion of observed phenotypic variation in the population examined can be attributed to genetic variation. It does not mean that the phenotype is fixed or cannot change (a form of genetic determinism).

An individual's DNA sequence may predispose them to certain phenotypes, but that predisposition does not fully determine who they are or how they will develop, because the environment can change or be manipulated to modify the phenotype. This is nicely illustrated with height, which has a heritability of ~80%.[242] Despite this very high heritability, there has been a well-documented increase in average height in many populations, likely due to changes in environmental conditions such as better nutrition and healthcare. Similarly, neuroscientific research has demonstrated that education (itself a form of population-level environmental manipulation) can increase cognitive abilities.[243] By extension, even if cognitive ability was 100% heritable, it is plausible that an environmental intervention could improve everyone's performance on a cognitive test but while everyone may improve with the intervention, the remaining differences between people would be for genetic reasons.

242 Yengo L, Vedantam S, Marouli E, *et al* (2022) A saturated map of common genetic variants associated with human height *Nature* **610**:704–12.

243 Ritchie SJ, Tucker-Drob EM (2018) How Much Does Education Improve Intelligence? A Meta-Analysis *Psychological Science* **29(8)**:1358-69.

Annex 4 Overview of a genome-wide association study (GWAS)

A genome-wide association study (GWAS) requires the collection of genotype and phenotype information for a large group of individuals. Typically, DNA is collected from research participants via saliva or blood samples, and the DNA is genotyped using SNP microarrays or DNA sequencing methods such as whole genome sequencing or whole exome sequencing (see **Annex 2**). SNP microarrays are the most widely used technology as they can genotype hundreds of thousands of SNPs spread throughout the genome quickly and cost-effectively.

SNP genotype data is accessible to researchers for some Biobanks, and for many large cohorts.

In a GWAS research participants are genotyped, and the resulting data undergo sample and variant quality control steps using established software tools and pipelines.[244] After this, genotypes are 'phased' which allows for the alleles of non-genotyped variants to be statistically inferred using reference populations. Genetically similar individuals are assigned to a **genetic ancestry** group, and relatedness of participants considered.

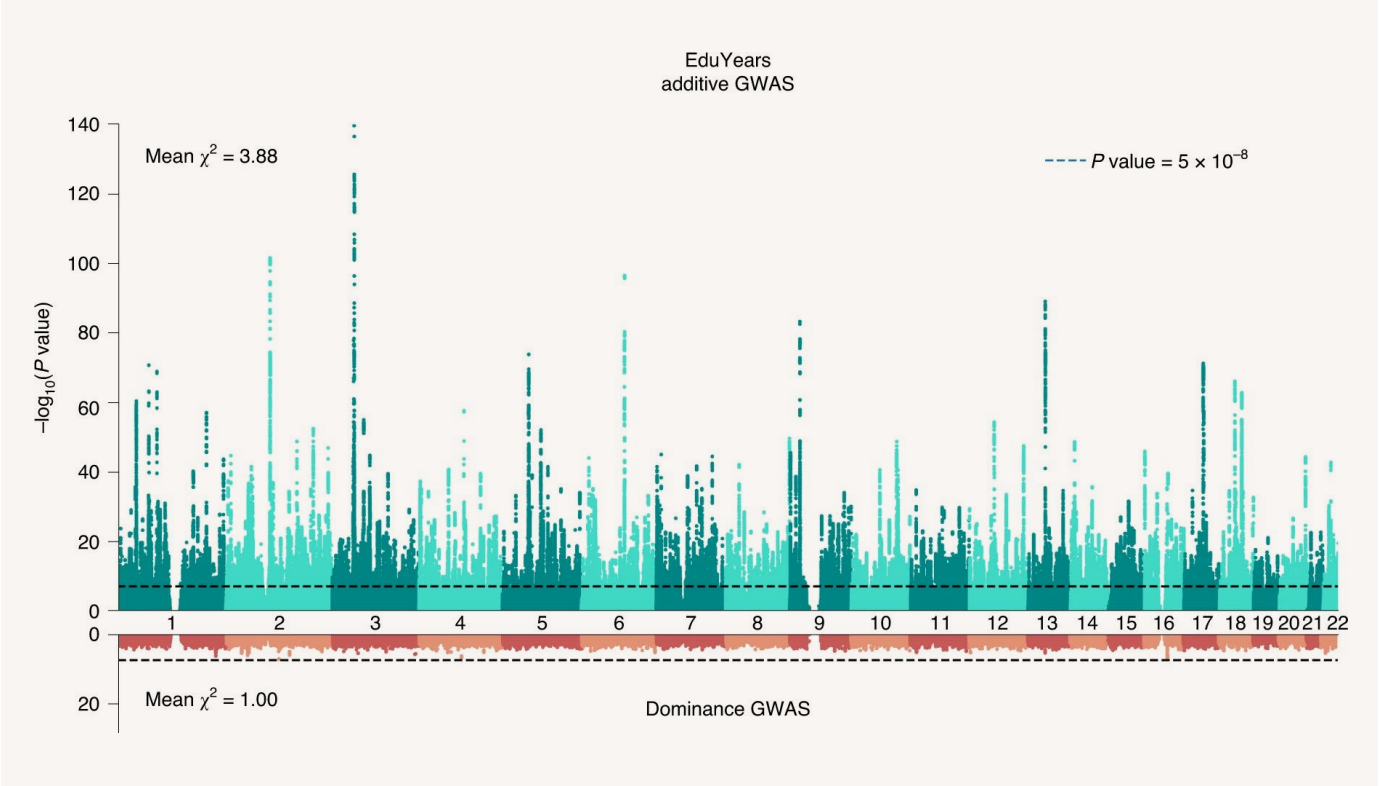
Once these steps are complete, the alleles of each variant (typically a SNP) are tested for association with the phenotype of interest using regression-based statistical methods. These methods typically assume additive effects of alleles on a phenotype. Each statistical test is performed on upwards of 600,000 genetic variants across the genome, resulting in many statistical tests. To account for this, a stringent **p-value** threshold of $5x10^{-8}$ is used to indicate statistical significance. Results are replicated in independent samples to ensure they are reliable.

Since the physical location of each genetic variant is known, a significant association points to a genomic region (or locus) that influences the phenotype. The results are visualised as a Manhattan plot, where tall 'skyscrapers' of signal indicate regions of the genome with evidence of genetic associations. See **Figure 8**.

244 Marees AT, de Kluiver H, Stringer S, *et al* (2018) A tutorial on conducting genome-wide association studies: Quality control and statistical analysis *International Journal of Methods in Psychiatric Research* **27(2)**:e1608.

Further *in silico* analyses are required to interpret these signals in a biological context, often by integrating epigenetic and transcriptomic data from other sources. Experimental approaches may also be used to test specific hypotheses.[245]

Figure 8: Manhattan plot of the most recent GWAS for years spent in education of ~3 million individuals using an additive (green) and dominant (red) statistical model. The x axis is the chromosomal position, and the y axis is the significance on a −log10 scale. The dashed line marks the threshold for genome-wide significance ($P = 5 \times 10^{-8}$).

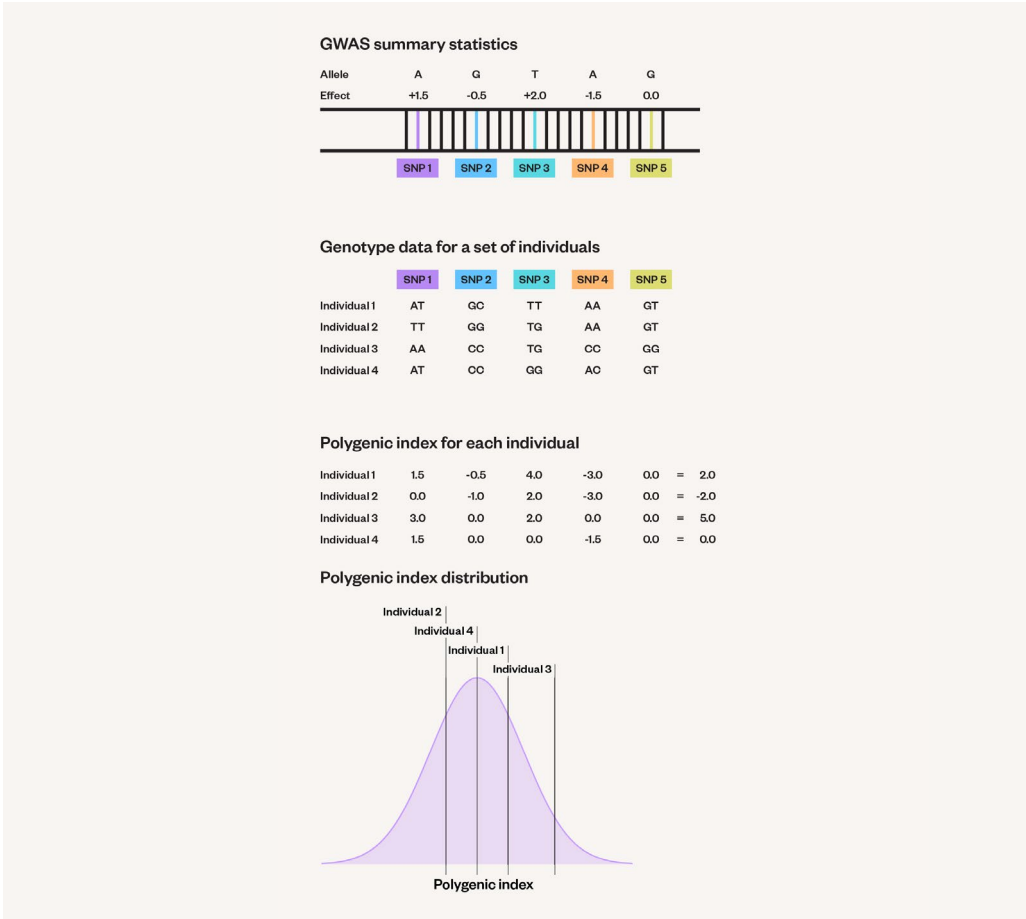


245 Uffelmann E, Huang QQ, Munung NS, *et al* (2021) Genome-wide association studies *Nature Reviews Methods Primers* 1:59.

Annex 5 Polygenic indices and how they are calculated

Polygenic indices (PGIs) are calculated as a weighted sum of alleles in a set of SNPs associated with a phenotype. The weights correspond to the direction (positive or negative) and strength (‘**effect size**’) of the association with the phenotype. Figure 9 illustrates how a PGI is calculated for four individuals using five SNPs. In practice, PGIs typically include thousands of SNPs.

Figure 9: A simplified example of how a PGI comprised of five SNPs is calculated for a given phenotype, for four individuals. PCIs are normally distributed in a population, meaning that most people will have a PGI close to the average, with only a small number of people having very high or very low PGI.



The creation of a PGI requires two datasets: (1) **GWAS summary data** for the phenotype of interest, which includes p-values and effect size for each SNP tested; and (2) genotype data for a set of individuals.

Various methods exist for calculating PGIs, differing in how SNPs are selected and how allele weights are calculated.[246] One commonly used approach is to rank SNPs in the GWAS data based on their association with the phenotype and select a subset that meets a certain threshold (e.g., SNPs with a p-value < 0.05). Once the SNPs have been selected, they are matched to the individual genotype dataset, ensuring both datasets contain the same set of SNPs.

For each individual, a genotype value is calculated for each SNP by counting the number of phenotype-associated alleles they carry (0, 1, or 2) and multiplying this value by the allele weight. For example, if allele 'A' for a given SNP increases the phenotype value by 1.5, an individual with two copies of the A allele would receive a value of 3 (2 × 1.5), while an individual with no copies of the A allele would get a value of 0 (0 × 1.5). A PGI for each individual is calculated as the sum of these genotype values across all selected SNPs. In **Figure 9**, the values for five SNPs are summed, resulting in a PGI of 2.0 for individual 1.

A key question is how many SNPs to include in a PGI. While PGIs can be created using only SNPs that are genome-wide associated with a phenotype (i.e. p value < $5x10^{-8}$), it is often the case that all SNPs examined in the GWAS are included – even though a high proportion will be given near-zero weights as they are not statistically associated with the phenotype of interest.

PGIs can provide a very rough estimate of an individual's genomic propensity for a phenotype relative to others in the same population, and can be used as a numerical variable in statistical analyses. It is important to note that PGIs do not capture all genomic influences on a phenotype, and changes in environmental conditions may alter PGI influences.

246 Ni G, Zeng J, Revez JA, *et al* (2021) A Comparison of Ten Polygenic Score Methods for Psychiatric Disorders Applied Across Multiple Cohorts *Biological Psychiatry* **1;90(9)**:611-20.

Glossary of key scientific terms used in the report

Adoption studies: In adoption studies, two types of pairs are examined: (1) individuals who are genetically related but do not share a common family environment (i.e. adoptees and their biological parents), and (2) individuals who share a family environment but who are not related (i.e. adoptees and their adoptive parents). By comparing how similar adoptees are to their biological versus their adoptive parents, researchers can separate out the relative importance of genetic and environmental influences on a specific phenotype. See also **twin studies** and **molecular genetic studies**.

Allele: An alternate version of DNA sequence (single base or series of bases) at a specific location in the genome. Where such **DNA variation** exists, an individual inherits two alleles – one from each parent – which may be similar or different.

Assortative mating: This occurs when individuals with similar phenotype values are more likely to pair and reproduce with each other than would be expected by chance. Assortative mating is found for educational attainment, intelligence and personality, with couples correlating positively for these phenotypes. It can lead to increased genetic similarity within couples, which can complicate the interpretation of genetic research findings. For example, if people with high levels of education tend to marry thin people, then the genetic variants influencing higher levels of attainment and lower body weight will become correlated in their offspring. As a result, a GWAS of attainment in the offspring may pick up associations for the (non-causal) variants for body weight, and vice versa.[247]

Candidate gene: A gene selected for study based on *a priori* knowledge of the gene's function or reported association with a phenotype. In general, candidate gene studies have been poorly replicated and have been superseded by **genome-wide association studies**.

247 Domingue BW, Fletcher J, Conley D, and Boardman JD (2014) Genetic and educational assortative mating among US adults *Proceedings in the National Academy of Sciences U.S.A.* **111(22)**: 7996–8000.

Chromosome: A single long molecule of DNA. In diploid cells (all cells of the body other than sperm and egg cells) the human genome is organised into 46 chromosomes: 23 pairs of autosomes (1–22) and one pair of sex chromosomes (typically XX = female; XY = male), with one member of each chromosomal pair inherited from each parent. Sperm and egg cells are haploid, containing only one copy of each chromosomal pair (23 chromosomes).

Complex (multifactorial) **phenotypes**: These arise as a result of a multitude of both genetic and environmental influences, and their interplay across the life course. This means that while the DNA sequence we inherit might influence how we think, feel and act, DNA alone does not determine who we are and how we develop.

Confounding: A confounder is a third variable, often unmeasured, that influences both the outcome and the risk factor, generating a spurious association between the two. This can lead to incorrect conclusions about cause and effect. **Gene–environment correlation** can be a source of confounding in observational research. For example, reading ability is genetically influenced, but offspring can both receive the genetic variants associated with reading ability and be influenced by a home reading environment. This is a form of passive gene–environment correlation, and it can confound observational associations between parental and offspring reading characteristics. Active and evocative gene–environment correlation can also create confounding if the environment that an individual experiences is influenced by their genotype.

DNA (deoxyribonucleic acid): The molecule that contains the genetic instructions for all living things. The DNA molecule is composed of two strands that coil around each other to form a double helix. Each strand contains a sugar-phosphate backbone, attached to which is one of four bases: adenine (A), guanine (G), thymine (T) or cytosine (C).

DNA sequence: This refers to the sequence of DNA bases (A, T, C, G) in the human genome. The order of bases is important because it encodes the biological information that cells use to develop and function.

DNA sequencing: The laboratory-based process of determining the exact order of nucleotides (or bases) in the DNA sequence. The sequence of bases forms instructions for an organism's cells to follow, such as how to create the proteins it needs to develop and function. DNA sequence can be determined for the whole genome (**whole-genome sequencing**; WGS) or just for the protein-coding regions (**whole-exome sequencing**; WES) of an individual.

Effect size: A statistic that quantifies the strength of a relationship between two measures.

Environment: In genomic research, 'environment' is taken to mean anything other than DNA sequence. Environmental factors examined in genetically informed educational research include family-, school- and neighbourhood-level factors. Environments that make two individuals in a family (such as siblings) similar are called 'shared', while those that make them dissimilar are called 'non-shared'.

Family-based GWAS: Family-based association tests that make use of first-degree relatives (typically, trio data comprised of both parents and a child). Family-based GWASs require larger sample sizes than GWASs of unrelated individuals to achieve the same statistical power but can help mitigate the influence of confounding factors.

Gene: Genes are arranged along chromosomes and consist of a sequence of DNA that is transcribed to produce a protein or other functional product. These products carry out biological functions inside or outside the cell or regulate the transcription of other genes. There are approximately 20,000 protein-encoding genes in the human genome.

Gene–environment correlation (rGE): Many environmental factors that relate to child development and functioning, such as parental discipline and life events, are heritable.[248] This indicates the presence of gene–environment correlation, where certain environments are more prevalent for individuals with certain genotypes, creating a link between their DNA and their experiences.[249] Gene–environment correlation can arise through three proposed mechanisms, passive, evocative and active, the first two of which are referred to in this report:

Active rGE: This occurs when children select their environments based on their genotype. For example, a child with a high PGI for attainment may join a maths club or hang out with the more academic kids in school.

Evocative rGE: This occurs when a child's genotype evokes specific responses from others, creating a correlation between that genotype and the environment. For example, a child with a high PGI for educational attainment (which may manifest in certain behaviours and cognitive traits) may evoke a specific response from teachers and parents, such as the creation of further learning opportunities.

Passive rGE: This can result from children inheriting both phenotype-associated genetic variants and phenotype-environments from their parents. For example, children of parents genetically predisposed to stay in education longer may themselves have higher educational attainment due to the direct effects on their own education of the variants associated with such attainment that they inherit. If these parents also create rearing environments correlated with their genetic propensity to education, the children inherit these educationally rich environments as well – such as more books in the home, extra tuition, or educational trips and experiences.

248 Kendler KS and Baker JH (2007) Genetic influences on measures of the environment: a systematic review *Psychological Medicine* **37(5)**: 615–26.

249 Plomin R (2014) Genotype–environment correlation in the era of DNA *Behaviour Genetics* **44(6)**: 629–38.

Gene–environment interaction (GxE): This refers to situations where vulnerability to an environmental exposure varies depending on the individual's genotype and, conversely, the effect of a genotype on a phenotype is modified by environmental exposures. When GxE interaction is present, a specific environmental change influences the phenotype of interest in different ways depending on the genotype.

Genetic ancestry: Ancestry is directly inferred from genomic variation data and not self-defined. It is used in genomics research to group genomes by how similar they are in patterns of genomic variation. Individuals are assigned to a genetic ancestry group as part of a GWAS analysis to avoid biased results owing to **population stratification**. The ancestry groupings are often given continental-level descriptors such as 'European', 'Asian' or 'African'. However, this approach overlooks the fact that ancestry is a continuum and not well captured by these population-level descriptors. Further, use of population-level descriptors that are linked to biology and can overlap with, or be conflated with, race and ethnicity can perpetuate harmful thinking and is both ethically and scientifically fraught. In response to these issues, the National Academies of Sciences, Engineering and Medicine (NASEM) recently released a report detailing a set of 13 recommendations on the use of population descriptors in genetics and genomics research.[250]

Genetic correlation: This measures the similarity (or correlation) between the genetic influences on two phenotypes. A genetic correlation of 0 indicates that genetic influences on one phenotype are independent of the other; a correlation of 1 indicates that genetic influences on both phenotypes are entirely shared.

Genetic nurture (also called dynastic effects): The genetic influences of family members on one another that operate via environmental pathways and processes.

Genetic variation (also called polymorphism): A position in the human DNA sequence at which there are at least two versions – or **alleles** – in the population. DNA variation comes in many forms: it may be a difference in a single base position (such as a single nucleotide polymorphism; SNP), a section of DNA that is deleted or repeated (such as copy number variations; CNVs), or the loss or duplication of entire chromosomes (such as trisomy 21). DNA variants are detectable at different frequencies in a population: some are common (more than 5% of individuals in the population have the alternate allele) while some are rare or ultra-rare (present in fewer than 0.1% of individuals). Some DNA variants will negatively impact biological function and lead to diagnosable conditions, while others have no obvious biological effect.

Genetics: The study of genes and how they function, as well as how they are inherited.

Genome: The complete set of DNA instructions present in a cell.

250 National Academies of Sciences, Engineering, and Medicine (2023) *Using population descriptors in genetics and genomics research: a new framework for an evolving field* (Washington, DC: National Academies Press).

Genome-wide association study (GWAS): The main research method used to identify genetic variants associated with heritable phenotypes. It involves comparing DNA variation data from a very large number of individuals that differ for the phenotype of interest (e.g. individuals with varying academic performance) to identify alleles that correlate with phenotype variation. GWASs can use a case-control study design when the phenotype of interest is dichotomous (e.g. cases with ADHD and controls without ADHD), or a quantitative approach when the phenotype is quantitative (e.g. intelligence). GWASs identify not necessarily the causal allele but rather a region of the genome that is correlated with the phenotype. These regions are followed up through further experiments to understand how the variation impacts biology.

Genomics: The study of the genomes of individuals and organisms that examines both the coding and non-coding regions. This term is also used when talking about related laboratory and bioinformatic techniques. The study of genomics in humans focuses on areas of the genome associated with health and disease.

Genotype: An individual's combination of **alleles** (or DNA variants). If an individual has two copies of the same allele for a DNA variant, they are said to be homozygous. If they have different alleles for the DNA variant, they are heterozygous.

GWAS summary data: The results of a GWAS, including a list of all tested genetic variants and their effect size. Summary data for thousands of GWASs have been made publicly available to download and query. The minimum required in a GWAS summary data file is a list of SNP IDs, SNP locations and genomic build, alleles, strand, effect size and standard error, p-value, test statistic, minor allele frequency, and sample size.

Heritability: This is a numerical value that estimates the degree to which variation in a phenotype (such as height or reading ability) in a population is due to genetic variation between individuals in that population. The more heritable a phenotype is, the more alike individuals who are also genetically similar will be for that phenotype. When heritability is estimated from twin studies, it is called 'twin heritability', and the remaining variation is assumed to be of environmental origin and classified as shared or non-shared. Heritability estimates from twin studies include all sources of genetic variation. When heritability is estimated using genetic variation data it is called 'SNP heritability' and is the variance explained by the additive effects of measured common DNA variants in a population. Heritability estimates are useful for guiding genomic research, for example to justify the search for more genetic variants using GWAS strategies, and it can provide limits on the performance of polygenic 'indices'.

Human genome: The full DNA sequence for an individual, consisting of 3.1×10^8 nucleotides, or bases. Diploid cells contain two copies of the genome, one inherited from each parent (to total 6.2×10^8 bases).

Linkage disequilibrium (LD): Correlations between nearby SNPs in the genome. LD patterns differ across genetic ancestries.

Mendelian (or **monogenic**) **condition:** A condition or disorder that arises due to a change in DNA sequence in a single gene. Under normal environmental conditions, the altered DNA sequence is both necessary and sufficient for the condition to occur. This contrasts with complex polygenic phenotypes that arise due to the aggregate effects of many DNA sequence variants (each of small effect) and environmental factors.

Mendelian randomisation: A method that uses genetic variants (usually SNPS) as 'instrument variables' to test for causal effects of environmental exposures on outcomes using observational data.

Missing heritability: The term given to the difference between twin-based heritability estimates for a given phenotype and the variance accounted for by genetic variants identified in a GWAS of the same phenotype. The missing component can be broken down into 'hidden' heritability, which is the difference between the genetic variants identified in a GWAS and the **SNP-based heritability** of the phenotype, and 'still-missing' heritability, which is the difference between SNP-based heritability and twin heritability. Still-missing heritability is thought to be due to other types of DNA variants that are not captured in current GWASs, and to gene–gene and gene–environment interaction effects. Hidden heritability can be recovered by increasing the sample size of the GWAS to find more genetic variants of (even smaller) effect.

Molecular genetic research: Rather than inferring genetic contributions to a phenotype by using individuals with varying degrees of relatedness (as in twin and adoption studies), DNA variation is studied directly and linked to measured differences between people for a phenotype. See also **adoption studies** and **twin studies**.

P-value: In GWASs, a p-value quantifies the probability of observing an association between a genetic variant and a phenotype by chance, assuming that no true relationship exists. A very small p-value (commonly less than 5×10^{-8}) suggests that the observed association is unlikely to be due to chance and the null hypothesis of no association is rejected.

Phenotype: Any measurable characteristic of an individual, for example a physical phenotype such as weight or height, cognitive phenotype such as intelligence, or behavioural phenotype such as aggression. Phenotypes may be classified as a diagnosable disorder (an individual either has conduct disorder or does not), or measured on a scale (individuals show different levels of aggressive behaviours). The term might be used interchangeably with 'trait'.

Pleiotropy: The phenomenon of a genetic variant influencing more than one phenotype. Broadly speaking, two main types of pleiotropy exist: (1) biological (also called horizontal) pleiotropy, where the DNA variant independently affects the two phenotypes, and (2) mediated (or vertical) pleiotropy, where the variant affects one phenotype, which in turn affects the second phenotype. When pleiotropy is observed it is difficult to figure out which scenario is the most likely without further research.

Polygenic index (PGI): A cumulative measure of an individual's genetic propensity for a specific phenotype based on the weighted sum of many thousands of DNA variants distributed throughout the genome. PGIs for diagnosable medical conditions, such as cancer, tend to be referred to as polygenic risk scores (PRS) or polygenic scores (PGS).

Polygenicity (adjective: **polygenic**): The contribution of many DNA variants (thousands or tens of thousands) to the variation in a phenotype.

Population stratification: This occurs when differences between populations in the frequency of genetic variants are correlated with environmental differences between populations by chance. If not properly accounted for, it can lead to spurious genetic associations in a GWAS. For example, if two populations are separated geographically, under certain conditions random genetic differences can accumulate over many generations. If a phenotype differs between these populations due to cultural, economic, social or broader environmental factors, it will appear in a GWAS as associated with the DNA variants that differ in frequency between the populations. If the patterns of genomic variation between the two populations are adjusted for, the association will disappear.

Population-based GWAS: A GWAS performed on a cohort where individuals are assumed to be randomly drawn from the population and unrelated. Typically, any pairs of relatives closer than second cousins are identified and removed prior to analysis.

Randomised controlled trial: A prospective, comparative, quantitative study/ experiment performed under controlled conditions with random allocation of intervention(s) to comparison groups. The randomised controlled trial is the most rigorous and robust research method of determining whether a cause–effect relation exists between an intervention and an outcome.

Single nucleotide polymorphism (SNP; pronounced 'snip'): A single-nucleotide (A, C, G or T) variation in the DNA sequence. SNPs are the most common type of genetic variation in human populations and usually consist of two alleles. They are easy to genotype at scale, making them the main type of genetic variation examined in GWASs. Not all SNPs impact gene function: a high proportion are located in the spaces between genes (i.e. in non-coding regions of the genome).

SNP microarray (also referred to as a SNP array or SNP chip): This is a laboratory technology used to genotype – determine the combination of alleles – an individual at many thousands of SNPs spread throughout the genome simultaneously. It is the most common technology used to generate genotype data for individuals analysed in a GWAS.

Statistical power: This is a measure of how likely a study is to detect a real effect or difference if one truly exists. It is influenced by factors such as the study design, **effect size**, variability in the data, and sample size (number of research participants). In a well-powered GWAS, researchers are less likely to miss true differences in the frequency of an allele (a version of a DNA variant) between individuals with a condition ('cases') and those without it ('controls').

Twin studies: These examine twins raised in the same environment but who differ in genetic relatedness, to tease apart shared environmental and genetic influences on a phenotype. Monozygotic (MZ, or identical) twins develop from a single egg fertilised by a single sperm and share all of their genes, whereas dizygotic (DZ, or fraternal) twins share on average half of their genes. Assuming that both sets of twins have equally similar family/rearing environments, finding that MZ twins are more similar for a phenotype than DZ twins indicates genetic influence. See also **adoption studies** and **molecular genetic studies**.

Whole-exome sequencing (WES): This involves sequencing only the protein-coding regions of the genome (around 2% of all DNA bases).

Whole-genome sequencing (WGS): This refers to DNA sequencing of the entire genome, including both coding and non-coding regions.

Nuffield Foundation
100 St John Street
London
EC1M 4EH

www.nuffieldfoundation.org
info@nuffieldfoundation.org
Nuffield Foundation

Nuffield Council on Bioethics
100 St John Street
London
EC1M 4EH

www.nuffieldbioethics.org
bioethics@nuffieldbioethics.org
Nuffield Council on Bioethics