Behavioral and Brain Sciences

cambridge.org/bbs

Challenging the utility of polygenic scores for social science: Environmental confounding, downward causation, and unknown biology

Callie H. Burt

Department of Criminal Justice & Criminology, Center for Research on Interpersonal Violence (CRIV), Georgia State University, Atlanta, GA, USA
cburt@gsu.edu; www.callieburt.org

Target Article

Cite this article: 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: 1–66. doi:10.1017/S0140525X22001145

Target Article Accepted: 27 April 2022
Target Article Manuscript Online: 13 May 2022
Commentaries Accepted: 27 August 2022

Keywords:
behavior genetics; environmental confounding; gene–environment correlation; genetic heterogeneity; GWAS; human potential; polygenic scores; population stratification; sociogenomics; statistical genetics

What is Open Peer Commentary? What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 20) and an Author's Response (p. 56). See bbsonline.org for more information.

Abstract

The sociogenomics revolution is upon us, we are told. Whether revolutionary or not, sociogenomics is poised to flourish given the ease of incorporating polygenic scores (or PGSs) as "genetic propensities" for complex traits into social science research. Pointing to evidence of ubiquitous heritability and the accessibility of genetic data, scholars have argued that social scientists not only have an opportunity but a duty to add PGSs to social science research. Social science research that ignores genetics is, some proponents argue, at best partial and likely scientifically flawed, misleading, and wasteful. Here, I challenge arguments about the value of genetics for social science and with it the claimed necessity of incorporating PGSs into social science models as measures of genetic influences. In so doing, I discuss the impracticability of distinguishing genetic influences from environmental influences because of non-causal gene–environment correlations, especially population stratification, familial confounding, and downward causation. I explain how environmental effects masquerade as genetic influences in PGSs, which undermines their raison d'être as measures of genetic propensity, especially for complex socially contingent behaviors that are the subject of sociogenomics. Additionally, I draw attention to the partial, unknown biology, while highlighting the persistence of an implicit, unavoidable reductionist genes versus environments approach. Leaving sociopolitical and ethical concerns aside, I argue that the potential scientific rewards of adding PGSs to social science are few and greatly overstated and the scientific costs, which include obscuring structural disadvantages and cultural influences, outweigh these meager benefits for most social science applications.

1. Introduction

Extraordinary techno-scientific advances over the past two decades have transformed human genetics. Scientists are now able to measure several million genetic variants across the genome (i.e., genome-wide) relatively cheaply (<$100) and efficiently with automated pipelines. Consequently, millions of individuals have been genotyped, which is the measurement of preselected variants, across the genome. Over the past decade, genome-wide association studies (GWASs), in which a phenotype (trait) is regressed on each of the millions of genetic variants with a few controls, have become the predominant method to statistically estimate genetic associations with genome-wide data and increasingly large datasets. Thousands of GWASs have been performed, identifying hundreds of thousands of significant associations with a multitude of traits and disease states (e.g., Buniello et al., 2019).

These molecular and computational innovations have launched the new science of sociogenomics, characterized by the application of cutting-edge statistical genetic tools and measures to social outcomes. In recent years, social scientists have teamed with biostatisticians and formed large consortia to conduct GWASs on complex social outcomes, such as educational attainment (Lee et al., 2018), same-sex sexual behavior (Ganna et al., 2019), number of children (Barban et al., 2016), and income (Hill et al., 2019), with large (and growing) genetic datasets. In sociogenomics, as elsewhere, GWAS results are commonly used to create genetic summary scores, known as polygenic scores (PGSs), representing the (additive) genetic propensity for some trait or behavior (e.g., years of educational attainment completed). Preconstructed PGSs have been incorporated into widely used social science datasets, such as the Add Health Study and Health and Retirement Study (HRS), to be dropped into models "just like any other variable," no genetic expertise required (Braudt, 2018). Given the availability and increased acceptance of genetics in social science, sociogenomics is poised to flourish.

This new "golden age" of sociogenomics filled the void left by the recent demise of the candidate gene × environment era, which was, by and large, a spectacular failure because of methodological limitations and an oversimplified biology (see Charney, 2022; Dick et al., 2015). Suggesting the candidate gene-era "should be a cautionary tale," psychiatric geneticist

© The Author(s), 2022. Published by Cambridge University Press

Matthew Keller asked: "How on Earth could we have spent 20 years and hundreds of millions of dollars studying *pure noise*?" (quoted in Yong, 2019; cited in Charney, 2022). With adjustments for multiple testing, attention to statistical power and large samples, and emphasis on replication, among other revisions, this nascent sociogenomics approach has addressed several methodological limitations plaguing the candidate gene approach. As a result, sociogenomics findings are touted as methodologically robust. Advocates are especially bullish about the potential of PGSs, which, they argue "just work" (i.e., are statistically significant genetic predictors) and have several potential social science applications that break through the stale, outdated nature versus nurture debate, on the one hand, and the neglect of genetics (or assumption of "genetic sameness") on the other (e.g., Belsky & Harden, 2019; Conley, 2016; Conley & Fletcher, 2017; Freese, 2018).

Further still, many sociogenomicists encourage other behavioral scientists to incorporate PGSs into their research (e.g., Braudt, 2018; Cesarini & Visscher, 2017; Harden, 2021b; Mills & Tropf, 2020). Pointing to evidence of ubiquitous heritability, the widening availability of genetic data, and the ease of incorporating PGSs into quantitative research, these scholars urge social scientists to incorporate genetics or risk losing out (e.g., Conley, 2016; Mills & Tropf, 2020). Others take an even stronger stance and emphasize not only the potential but also the necessity of incorporating genetics into social science, arguing that social science research that neglects genetics is, at best, partial and potentially flawed and misleading (e.g., Braudt, 2018; Harden, 2021a; Hart, Little, & van Bergen, 2021; Kweon et al., 2020). In her recent book, *The Genetic Lottery*, Harden (2021a) contends that social science sans genetics wastes time, resources, attention, and effort; supports misguided models of human behavior; and misinforms policies, causing still further damage. This neglect of unmeasured genetic heterogeneity makes social science research vulnerable to sweeping dismissals from other scientists (Freese, 2008) or political extremists (Harden, 2021a).

Yet it remains the case that only a paucity of behavioral science research includes genetics. This "neglect of genetics" is, some proponents have argued, not because of valid scientific reasons but of an ideologically motivated "tacit collusion" to ignore genetic differences between people among social scientists (Freese, 2018; Harden, 2021a; Wright & Cullen, 2012). Harden (2021b) argues that this alleged tacit collusion is not just misguided or morally "wrong in the way that jaywalking is wrong" but, given the scientific warrant to include genetics, it is "wrong in the way that robbing banks is wrong." Harden avers that "Failing to take genetics seriously is a scientific practice that pervasively undermines our stated goal of understanding society so that we can improve it" (p. 186). On this view, if progressive social scientists really want to ameliorate inequality, they need to get with the science and add genetics to their research.

Here, I scrutinize proponents' arguments about the significant value of PGSs for social science and with it the need to incorporate genetics into social science models. I do so not by questioning the ethical or sociopolitical implications of this work, as is common, but by scrutinizing the science of sociogenomics. Specifically, I focus on the utility of PGSs for social science and the key premises underlying their use as measures of "genetic propensities" for behavioral differences. Drawing on contemporary statistical genetic research, I explain how methodological limitations produce environmentally confounded PGSs. I emphasize that environmentally confounded genetic associations with complex social outcomes is not simply a tractable empirical problem to be addressed with more sophisticated methods. Rather, such confounding is inevitable when attempting to map layered and contingent social behaviors, like educational attainment, to a score representing a linear summation of base-pair differences, which themselves represent an entirely different set of layered contingencies. I explain why this inevitable environmental confounding of PGSs for complex social traits undermines their use as "genetic influences on" or "genetic potential for" social traits and achievements – as is common. After outlining the limitations of current sociogenomics methodologies, I consider the practical implications by examining several existing applications of PGSs to social science and their substantive contributions.

My explicit aim is to challenge the claim that genomics has much to offer social science, so much so that social science sans genetics is fatally flawed, scientifically indefensible, and possibly even morally suspect. I argue that, leaving sociopolitical risks aside, the potential scientific rewards are few and greatly overstated, and the potential scientific costs – obscuring environmental influences, perpetuating a flawed concept of genetic potential for social behaviors and achievements, and wasting resources – outweigh these meager benefits for most applications. I am not alone in my concerns, and not all sociogenomics practitioners are sold on the touted benefits of PGSs; however, cautious and skeptical arguments are invariably drowned out by enthusiastic hype and promissory notes. Much of the excitement around sociogenomics comes from the application of these new measures and techniques without clearly acknowledging limitations or accounting for well-known biases. Given this situation, my goal is to draw attention to and explicate the limitations of sociogenomics methods, especially PGSs, that vitiate their utility in the behavioral sciences.

Before moving forward, a few remarks about the larger backdrop are in order. Most historical and current critiques of social science genetics emphasize sociopolitical or ethical considerations rather than scientific concerns. This focus is because of both socio-historical reasons (racist and eugenicist applications and/or interpretations of this work in the past) and the fact that the advanced biology and statistical genetic methods of sociogenomics are well outside the bailiwick of most social scientists (and thus lack of expertise and skills to critically engage with this research). Here, I do not concentrate on sociopolitical or ethical concerns about sociogenomics research, because existing scholarship addresses these issues, acknowledging historical misuses with some atrocious results and highlighting the potential misrepresentation of sociogenomics findings to support genetic determinist and inferiorizing claims (e.g., Bliss, 2018; Duster, 2015; Harden, 2021a; Herd, Mills, & Dowd, 2021; Martschenko,

Callie H. Burt, a sociologist, is Associate Professor of Criminology at Georgia State University (GSU) and Faculty Affiliate of the Center for Research on Interpersonal Violence (CRIV) at GSU and the Center for Studies in Demography and Ecology (CSDE) at the University of Washington. Much of her work focuses on understanding the individual psychosocial mechanisms linking social disadvantages to an increased risk of criminal and health-risk behaviors. With support from a K01 fellowship from the Eunice Kennedy Shriver National Institute for Child Health and Human Development, Burt has spent the past several years training in statistical genetics and genomics.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Trejo, & Domingue, 2019). Although I share these concerns, my current focus is scrutinizing sociogenomics with the aim of fostering a dialogue that focuses squarely on the science.

This critical analysis proceeds in several parts. First, I provide a brief overview of the genetic and statistical genetic fundamentals necessary to understand these models and their limitations, recognizing that sometimes social scientists' lack of expertise in genetics and statistical genetics methods is a key barrier to engagement. (Readers wholly unfamiliar with genetic concepts can see the primer in Appendix A, whereas those familiar with sociogenomics concepts and methodologies may opt to jump to sect. 4.) Next, I describe proponents' key arguments for the value of adding genetics to social science. I then discuss and critique the key premises underlying these arguments, with a particular focus on explicating intractable environmental confounding in GWAS associations and PGSs.[1] I then explain how these challenges undermine the utility of PGSs as measures of genetic influences or potential. I conclude by offering several suggestions for the field.

2. A primer on genomics

At present preconstructed polygenic scores (PGSs) are available in several accessible social science datasets available to be dropped into models just like any other variable (Braudt, 2018; Mills & Tropf, 2020). Properly interpreting the meaning and challenges of PGSs, however, requires some knowledge of what PGSs capture, what they don't, and what these models assume.

2.1 Basic genetic concepts in sociogenomics

See Table 1.

2.2 Genetic variants, function, and prevalence

Given that sociogenomics focuses on genetic variation among people, understanding the type, prevalence, and distribution of human variation is necessary to understand what is and is not being captured in these studies. Genetic variants can be classified into three types: (1) single-nucleotide variants (SNVs), which are single-base changes (G → A); (2) indels, which are insertions of base pairs or deletions up to 50 bp and often involve tandem repeating units (e.g., GATA repeated 2–8 times); and, (3) structural variants (SVs), which are DNA rearrangements (deletions, duplications, or inversions) ranging from 50 bp to more than a million base pairs (1 Mbp). As discussed below, GWASs and PGSs analyze a subset of "common" single-nucleotide variants, known as single-nucleotide polymorphisms (SNPs), where common usually means present in at least 1% of the population (see Appendix A for more details).

Human genetic variation is extensive – all genetic variants compatible with life are likely represented in some individual living today (McClellan & King, 2010). Comparing the genomes of any two humans around the world, we would typically find between 3 and 4.5 million genetic differences between them or approximately 1 variant every 800 bases.[2] Most of these genetic variants are SNPs and are non-functional. That is, they have no effects on biological functioning or differences between people. Obviously, only functional variants contribute to differences between people. Although some genetic variation is debilitating, most genetic variation in a given genome is benign, ancient, and common.

In contrast, functional variants are those that either alter gene product (the protein produced) or gene dosage (e.g., the amount of protein produced). As an example of the former, the *SLC24A5* gene encodes a protein involved in epidermal melanogenesis and skin pigmentation through its intracellular potassium-dependent exchanger activity (Ginger et al., 2008). Several thousand years ago, a G → A mutation in *SLC24A5* occurred among people migrating from Africa to Europe. This variant, which changes the encoded amino acid from alanine to threonine, disrupts melanogenesis and thereby results in lighter skin tone (Lamason et al., 2005). Other variants can affect function not by changing the protein produced but, for example, by affecting the binding sites for various RNAs in a manner that reduces or increases transcription and thereby contributes to trait differences by altering gene dosage (the production of too much or too little of the functional protein).

All three variant types can be functional and contribute to differences between people. Although rare compared to SNVs and indels, evidence suggests that SVs have a disproportionate role in shaping human differences compared to other variants (Chiang et al., 2017; Collins et al., 2020; Takumi & Tamada, 2018). SVs can involve multiple copies of genes or the deletion of a gene and thus influence gene dosage. Sudmant et al. (2015) estimated that SVs were 50 times more likely than SNVs to affect gene expression and three times more likely to be associated with a trait difference than an SNV.

Despite being the extreme minority among the variants we carry, we all have thousands of functional variants in our genomes. A recent deep sequencing study of diverse ancestries identified approximately 11,700 functional variants per individual genome (Taliun et al., 2021). Another study of roughly half a million people in the United Kingdom, Backman et al. (2021) observed an average of ~600 variants, including 50 putative loss-of-function (pLOFs) variants, per gene. Backman et al. (2021) estimated that on average each of us carries 214 pLOF variants as "defective" gene copies. Although this variation is nontrivial, recall that we receive two copies of our genes (excepting the male-specific genes on the Y chromosome). In addition, a host of cellular mechanisms, including those shaping gene expression, compensate for many of these loss-of-function variants and facilitate robustness to functional mutations by, for example, up-regulating transcription (thereby producing more mRNA transcripts) and slowing the rate of mRNA decay (thereby increasing the ability of the cell to generate more polypeptides from the same mRNA transcript) (see Strachan & Read, 2018).

In addition to the several million genetic variants passed down by each of our parents, we inherit roughly 30–80 new mutations that arise during meiosis. The human population explosion over the past several hundred years has produced an abundance of new mutations as rare variants. Rare variants are disproportionately deleterious. Fu et al. (2013) estimated that ~86% of all deleterious SNVs are rare and recent. Many of these variants are found in only a handful of related people and are not represented in population samples. As discussed later, despite their prevalence and disease-relevance, rare variants pose a challenge for GWASs.

2.3 A brief note on ancestry and continental populations

Most sociogenomics studies at least briefly discuss ancestry and issues related thereto. A basic understanding of what this refers to is helpful (for a social science discussion, see Herd et al.,

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Table 1. Glossary, acronyms, and definitions for sociogenomics terms

Concept	Acronym	Definition
Allele		A version or alternative form of a DNA sequence (e.g., a version of an SNP) or a gene.
Allele frequency		The proportion of all variants at a given position that are the specific allele in question; usually reported as the frequency of the second most common variant (i.e., "minor allele frequency").
Copy number variant	CNV	A type of genetic variant in which the number of copies of a particular sequence varies between individuals.*
Gene		Sequences of DNA interspersed at irregular intervals on our chromosomes that serve as templates for making an RNA product.
Genetic risk score	GRS	Alternative for PGS
Genome		The total DNA sequence in an organism or cell; the human genome consists of roughly 6 billion nucleotide bases of nuclear DNA separated into 46 chromosomes plus mitochondrial DNA.
Genome-wide association study	GWAS	A statistical analysis that estimates the partial correlation between each measured DNA variant (usually SNPs) and a particular phenotype, net of a few controls (usually age, sex, and ancestry PCs).
Haplotype		A sequence of alleles found at linked loci on a chromosome.
Haplotype block		Blocks of variants that are in linkage with each other but not with variants in adjacent blocks (separated by recombination); the consequence of shared ancestry.
(Short) insertion–deletion	Indel	Narrowly, a change where one or more nucleotides are inserted (or deleted) in a sequence; broadly as used here, all types of DNA change that cause a size change at a specific position: insertions, duplications, deletions, and compound insertion/deletion up to 50 bp (includes short CNVs).
Linkage disequilibrium	LD	When particular alleles at separate loci are associated with each other at a significantly higher frequency than would be expected by chance.
Locus		Designated region on a chromosome. Can refer to a single-base position or a broader region.
Non-coding RNA	ncRNA	RNA that does not code for a protein; ncRNA has many functions in the cell.
(Genetic) principal components	PCs	Orthogonal controls for ancestry created from a principal components analysis (a dimension reduction technique) of the genetic relatedness or allele dosage matrix.
Polygenic index	PGI	Alternative for PGS
Polygenic risk score	PRS	Alternative for PGS
Polygenic score	PGS	A genetic summary score representing the additive genetic association with a trait; composite measure created as the sum of the GWAS-weighted allele dosages for each individual; human equivalent to the breeding value.
Quantitative trait locus	QTL	A locus (that statistical analysis has) linked to a continuous (quantitative) trait, like height.
Single-nucleotide polymorphism	SNP	A position on the genome where two (or more) alternative nucleotides are common (>1%) in the population; common SNVs.
Single-nucleotide variant	SNV	A position on the genome where alternative nucleotides exist.
Structural variant	SV	Sequence changes (insertions, deletions, translocations) that involve a change in more than 50 bases. (In the past, structural variation was concerned with large sequence changes of >1 kb, but with next-generation sequencing, SVs, it has come to represent smaller changes.)
Tag SNPs		Mostly non-functional SNPs in GWASs used to tag a region of common variation; common SNPs used to tag haplotypes.

*The term copy number variant used to be applied to all variants that had a variable number of tandem repeats, including short tandem repeats, such as the microsatellite in (D) where there are 12 or 11 copies of the CA dinucleotide. In genome sequencing projects, the term is reserved for large size changes only, such as variable numbers of repeats exceeding 50 nucleotides in the case of the 1000 Genomes Project. (Strachan & Read, 2018)

2021). Modern humans are, of course, a single species, which emerged some 550–750 thousand years ago (Fu et al., 2016). Although terminology varies, several population genetic studies classify humans roughly into five continental populations: African (AFR), European (EUR), East Asian (EAS), South Asian (SAS), and American (AMR), differentiated by their continental migration out of Africa within the last 100,000 years (The 1000 Genomes Project Consortium, 2015). Importantly, these populations are abstractions from an underlying continuum of genetic relatedness and *should not* be thought of as genetically distinct subpopulations (Coop & Przeworski, 2022; Feldman, Lewontin, & King, 2003).

The vast majority of variants in an individual's genome are shared by all continental populations (The 1000 Genomes Project Consortium, 2015). Only a small proportion of the variants in an individual genome are restricted to one continental population, and these tend to be recent mutations that are also rare in the populations in which they are found. However, allele frequencies for common variants do differ across groups because of population patterns of migration and mating, shaped by physical boundaries and sociocultural influences. Furthermore, allele frequencies vary in a more fine-grained manner across subgroups of populations, especially for rare variants (Mathieson & Mcvean, 2012). As discussed later, this variation in mostly random allele

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

frequencies across difference groups poses a major challenge for GWASs by inducing or inflating genetic associations through confounding between genotypes and outcomes (e.g., Berg et al., 2019; Morris, Davies, Hemani, & Smith, 2020a).

3. Statistical genetic methods of sociogenomics

3.1 What genetic differences are measured?

The complexity of GWASs/PGSs and the way that they are discussed can produce confusion over what is measured in these studies. Readers can be excused from thinking that these studies measure genes and/or causal variants that shape differences through some known biological pathway. The abstract of a recent study, for example, referenced "mothers with more education-related genes" (Armstrong-Carter et al., 2020). Genes are not measured in these studies. Rather, these studies measure and analyze a select subset of one form of variation in the genome: Single-nucleotide polymorphisms (SNPs) that have two alleles (e.g., A or C) (see Appendix A for a detailed discussion).[3] In this section, I describe with as much simplicity as possible what is measured in GWASs/PGSs and why. Although intricate, understanding what GWASs/PGSs do measure (SNPs) and that they do not measure (genes or causal variants) is necessary to understand the inherent limitations with this approach.

The GWAS methodology is rooted in the blocklike structure of our genome. Although technical detail is out of scope, we inherit whole chromosomes from each parent, but these chromosomes are composed of unique blends of blocks of our parents' maternal and paternal chromosomes created during the process of "crossing over" (or genetic recombination). Each chromosome we inherit is a unique blend of our parents' matching chromosomes, created when segments are exchanged in meiosis (an average of 1.5 blocks of exchange per chromosome). Helpfully, crossing over does not occur randomly across the genome but tends to occur in 1–2 kb regions, known as recombination hotspots, which occur every 50–100 kb across the genome (Myers, Bottolo, Freeman, McVean, & Donnelly, 2005). Consequently, blocks of chromosomal segments are passed down across many generations unbroken by recombination, and, by dint of being passed down unbroken, contain correlated SNPs (i.e., SNPs that are not inherited independently). These chromosomal segments that exist between recombination hotspots are known as *haplotype blocks*. The association between SNPs on a haplotype is known as *linkage disequilibrium* (LD) and exists as a matter of degree (as a correlation).

This haplotype structure of our genome means that there is much less variability between genomes than would occur from the random assortment of SNPs. For example, the average haplotype block contains ~50 SNPs, which would, in theory, allow 2^{50} different combinations. Typically, however, most haplotype (>90%) blocks will be characterized by six or fewer combinations of alleles (The International HapMap Consortium, 2005). The combination of alleles on a haplotype block is known as haplotype and represents ancestral segments defined by common, ancient SNPs. Rarer variation exists as heterogeneity around the SNPs that define haplotype blocks (Strachan & Read, 2018).

This haplotype structure of our genomes undergirds the GWAS methodology. Measuring and testing each of our 3 bn base pairs is impracticable. Instead, GWASs analyze a smaller number of SNPs from across the genome to tag regions of common variation (i.e., haplotypes). Contemporary GWASs scan the genome for associations between several millions of these preselected SNPs, known as "tag SNPs" and a trait. Significant SNP associations mark a genomic region ("genomic risk locus" or quantitative trait locus, QTL) in which an unknown causal variant(s) driving the association is presumed to lie. Tag SNPs are thus usually non-functional, common variants used as proxies for some unknown causal variant(s) in proximity (with which they are in LD). Proximity is relative and varying. Genomic risk loci can range in size from several hundred thousands to more than 1 Mbp.

Crucially, rare and more likely deleterious variants are not well tagged by SNPs, given that SNPs tag haplotypes defined by shared common variants, and most haplotypes will not contain the rare variants (or they wouldn't be rare)[4] (Backman et al., 2021; McClellan & King, 2010; Tam et al., 2019). Additionally, other variant forms – indels, copy number variants (CNVs), and SVs – are not measured in GWASs, and many are not well-tagged by common SNPs (Backman et al., 2021; Tam et al., 2019).

Additionally, because different ancestral groups can have different allele frequencies, different patterns of LD, and somewhat different haplotypes, tag SNPs often do not work in the same way across populations, even when the causal variant is the same (Martin et al., 2017; Peterson et al., 2019). This ancestral variation in LD and haplotypes is one biological reason why GWAS findings do not "port well" or generalize across ancestral groups (e.g., Mostafavi et al., 2020).

The haplotype structure of our genome also enables GWASs by facilitating imputation. GWASs rely on large samples; however, studies vary in the genotyping platforms they use, which measure somewhat different SNPs, and contain missing data. Knowledge of haplotypes allows the probabilistic imputation of missing or untyped genotypes at adjacent SNPs using more densely genotyped samples or whole-genome reference panels.[5] Most genotype arrays now measure between 500,000 and 2 million SNPs, and most contemporary GWASs now include ~10 million SNPs, most imputed (Tam et al., 2019).

The original aim of GWASs was to understand the underlying molecular basis of trait variation by tracing causal pathways from genetic variants to outcomes. The idea was that tag SNPs could be used to mark risk loci that could be followed up with fine-mapping and functional annotation to identify causal variants in genes with well-defined functions. Although GWASs have, in some cases, facilitated the identification of causal variants involved in disease pathogenesis, for reasons that are out of scope, biological interpretation is exceedingly difficult, in general, and even more so for complex social traits with increasingly numerous (>1,000) GWAS hits and miniscule effect sizes (see e.g., Backman et al., 2021; Crouch & Bodmer, 2020; Edwards, Beesley, French, & Dunning, 2013). Hence, sociogenomicists primarily use GWAS results for polygenic prediction, explicitly deemphasizing inquiry into causal variant(s) or biological pathways (but not always, see e.g., Ganna et al., 2019). In what follows, I briefly describe the nuts and bolts of GWASs, this is followed by a discussion of PGS generation.

3.2 GWAS methodology

GWASs are a theory-free analytic approach to scan the genome for trait-associated tag SNPs. This involves testing, for each SNP one at a time, whether an allele (SNP variant) is more common in cases versus controls (or for continuous traits, across different levels). GWASs thus test for independence between genotype and

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

outcome for each SNP, with a few controls (not including other SNPs). The 2018 educational attainment GWASs, for example, assessed whether allele frequencies – for roughly 10 million SNPs – differed (across groups stratified by) years of education (Lee et al., 2018). Typically, the type of effect of interest in GWASs is variant substitution effects, which can be understood as the counterfactual change in an individual outcome that would occur from changing the individual's genotype for a particular SNP at conception (holding all else constant) (Freese, 2008; Morris et al., 2020a). This counterfactual model assumes that genetic associations indicate a causal path from an individual's genotype (or allele dosage) to complex social traits, reflecting a variant substitution effect (Lawson et al., 2020).

The basic form of GWASs is straightforward. Here I focus on these basics, including the familiar linear equation form that underlies the model. This model has been elaborated in recent years, but the underlying logic remains the same. Using bi-allelic SNPs and assuming additive SNP effects, genotypes for a particular SNP (e.g., AA, AC, CC) are translated to numeric *allele dosage effects* by counting the number of minor (or effect) alleles (0, 1, or 2) for each individual. Allele dosages for each SNP are the focal independent variable in each of these millions of regressions (again, one for each SNP examined separately), which take the following general form:

$$Y = \beta_0 + \beta_1 \times \text{SNP} + \beta_2 \times \text{Sex} + \beta_3 \times \text{Age} + \beta_4 \times PC1 \ldots \beta_{14} \times PC10 + e$$

where Y is a continuous variable (e.g., years of education), and SNP represents the allele dosage measure, controlling for age, sex, and usually 10–20 genetic ancestry principal components (PCs, discussed shortly). The outcome of interest in this model is β_1 – the effect size for each SNP – which can be interpreted as the marginal effect of having one more minor allele (a unit increase in allele dosage) and its associated p-value. For binary outcomes, this would just approximate the form of a familiar logistic regression model. These results for the millions of separate regressions are automatically compiled into results by modern computational programming software, such as PLINK (Purcell et al., 2007) and METAL (Willer, Li, & Abecasis, 2010). Focal GWAS results, as the SNP effect size estimates and p-values, are known as *summary statistics*, which provide the input for further analyses. Summary statistics are often considered the "data" in GWASs even as these are more accurately referred to as the results (of the first step of the analysis) (Burt & Munafò, 2021).

Following the estimation of the GWASs from the primary study sample or the "discovery" sample, a number of diagnostic tests (e.g., Manhattan and QQ-plots, which display p-values on a $-\log_{10}$ scale) are performed (see Choi, Mak, & O'Reilly, 2020; Schaid, Chen, & Larson, 2018). Because of LD (non-independence among SNPs sharing a haplotype) and the examination of each SNP separately, there will invariably be multiple (even dozens of) SNPs marking a risk locus. Thus, follow-up analyses (e.g., clumping and thresholding) are conducted to define clusters of SNPs in high LD (often high LD is defined as $r^2 > 0.1$[6]) and to identify a single "lead SNP," usually the SNP with the lowest p-value, to represent this clump and mark a risk locus. In this way, risk loci (or QTLs) are defined as trait-associated regions marked by approximately independent ("lead") SNPs.

As noted, risk loci range in size from ~50 kbp to over 1 Mbp (e.g., Lee et al., 2018). Thus, a GWAS that reports 1,237 lead SNPs can thus be understood as identifying 1,237 approximately independent risk loci defined by a lead SNP and in which the causal variant(s) responsible for the association is presumed to lie. Such risk loci, which often stretch across multiple haplotypes, usually contain thousands of SNVs along with SVs and indels, and often multiple genes (hence the difficulty of biological interpretation).

Importantly, lead SNPs for complex social traits are invariably very weakly associated with an outcome, usually accounting for less than 0.01% of the variation. In their educational attainment study, for example, Lee et al. (2018) reported that "the median effect size of the lead SNPs corresponds to 1.7 weeks of schooling per allele." Similarly, among the five lead SNPs identified in their study of "non-heterosexuality" in the UK Biobank, Ganna et al. (2019) observed "very small effects"; "males with a GT genotype at the rs34730029 locus had 0.4% higher prevalence of same-sex sexual behavior than those with a TT genotype (4.0 vs. 3.6%)" (p. 3). Given the impracticability of biological interpretation and the weak prediction from any single variant or QTL, researchers have shifted to creating genetic summary scores that aggregate SNPs weighted by their effect sizes, discussed next.

3.3 Polygenic score (PGS) construction

Calculating PGSs (also called polygenic risk scores [PRSs] or genetic risk scores [GRSs], usually when referring to adverse biomedical outcomes) is now a common application of GWASs to predict complex traits (or disease risk) from weight and height to depression and educational attainment (Evans, Visscher, & Wray, 2009; Wray, Goddard, & Visscher, 2007). PGSs operate under a massively polygenic, additive model (Boyle, Li, & Pritchard, 2017). Under this model, summing the GWAS-weighted risk (or effect) allele dosages (0, 1, or 2) usually with several sophisticated statistical adjustments can provide an index of a continuous underlying (additive) genetic liability for a trait.[7] The human equivalent of the "breeding value" is in selective plant and animal breeding in human populations (Meuwissen, Hayes, & Goddard, 2001), and PGSs have been described as "summariz[ing] the cumulative effects of many variants across the genome and aim[ing] to index an individual's genetic liability for a given trait" (Domingue, Trejo, Armstrong-Carter, & Tucker-Drob, 2020, p. 465) or a "single quantitative measure of genetic predisposition" (Mills, Barban, & Tropf, 2018). The educational attainment PGS has been characterized as measuring "an individual's genetic predisposition for completing [more years of] formal schooling" (Bolyard & Savelyev, 2020) and a "DNA-based indicator[] of propensity to succeed in education" (Harden et al., 2020).

The specific details on PGS construction can, and have, filled articles (see Choi et al. [2020] for more details), but the basic process is as follows: Run GWAS in discovery sample → replicate results in an independent sample → adjust for LD using a reference panel → select SNPs → adjust for LD and winner's curse → construct PGS → test PGS prediction in a target sample → assess PGS with incremental R^2. It is worth noting that there are several decisions by researchers involved in PGS construction. In the "select SNPs" phase of PGS construction, researchers decide which SNPs to include in the PGS (via p-value thresholds) based on the success of prediction. Specifically, researchers evaluate several PGSs created at a variety of p-value thresholds and select the best PGS predictor (measured by R^2), which is usually

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

the PGSs created from a $p < 1$ threshold (i.e., *no* p-value threshold) (e.g., Belsky et al., 2018; Ganna et al., 2019; Lee et al., 2018).[8]

Thus, in what may come as a surprise to some, most PGSs are constructed from all available SNPs regardless of their statistical significance in the GWAS. Available evidence suggests that these "all SNPs" PGSs are more environmentally confounded than those that use (more stringent) p-value thresholds, such that while these may explain more variance, they do so because they capture environmental influences as well as genetic ones (Berg et al., 2019; Mostafavi et al., 2020).

4. The utility of PGSs for social science: Proponents' arguments

Touted as a powerful new "tool" for social scientists to incorporate genetics into their research, PGSs are said to offer exciting new opportunities for social science research (Braudt, 2018; Freese, 2018; Harden & Koellinger, 2020; Mills & Tropf, 2020). Below I describe proponents' chief arguments about the utility of PGSs for social science, but first a note on PGSs lack of efficacy in individual prediction.

With few exceptions (e.g., Plomin, 2019; Plomin & Von Stumm, 2018), scholars agree that PGSs do not predict complex social outcomes with any degree of efficacy or accuracy and, therefore, should not be used for individual prediction (see, e.g., Harden & Koellinger, 2020; Morris, Davies, & Smith, 2020b). Although not appropriate for predicting individual outcomes, proponents emphasize myriad of benefits to incorporating PGSs to social science.

4.1 "Getting genetics out of the way"

Perhaps the most hyped value of PGSs in social science is to control for genetic heterogeneity in studies of environmental effects. According to Harden (2021a), many sociogenomicists are most excited about the potential of PGSs as a tool "to make genetics recede into the background, to get it out of the way" so that we can more clearly see the effects of environments (see also Conley, 2016). Given ubiquitous heritability, proponents argue that uncontrolled genetic heterogeneity poses a serious threat to inferences about the effects of specific environments, as these ostensibly environmental causes may be biased or spurious (as actually driven by genetic differences) (Harden & Koellinger, 2020; Hart et al., 2021). For example, rather than health or longevity being influenced by higher educational attainment, scholars have suggested, these relationships may be spurious with genetic endowment being the causal force. Similarly, sociogenomicists have asked, whether parental environments, including early childcare, causally influence educational attainment or whether these are spuriously associated because of shared genetic endowments.

Proponents also argue that incorporating PGSs as control variables into social science research can enhance the precision of environmental estimates (Cesarini & Visscher, 2017; Harden, 2021a, 2021b; Kweon et al., 2020). This enhanced precision may increase the power associated with randomized controlled trials, potentially shrinking their cost (Lee et al., 2018; Rietveld et al., 2013). Controlling for genetic heterogeneity with PGSs, proponents argue, may also reveal previously obscured environmental effects. For example, some environmental influences on educational attainment may only be apparent among those at "high genetic risk" (Herd et al., 2021). For these reasons, proponents suggest, PGSs are valuable as a control for differential genetic propensity to illuminate more clearly and precisely the effects of environmental influences (Harden & Koellinger, 2020; Trejo & Domingue, 2019).

4.2 A powerful, flexible analytic tool for causal inference

Proponents also emphasize the value of PGSs as a powerful tool for causal inference (Belsky & Israel, 2014). This strength of PGSs, proponents argue, draws on several unique advantages of genetic data (Conley, 2016; Harden, 2021a). First, evidence (from twin studies of heritability) suggests that genetic differences matter. Second, "the genetic sequence of each person is fixed at conception and does not change throughout one's lifetime" (Kweon et al., 2020), which means that genotype need only be measured once. Further, once measured, PGSs can be calculated for any outcome, which need not be measured in the study, and as PGSs are updated with larger and more diverse samples, these individual scores can be created and updated (Belsky et al., 2018; Harden, 2021a, 2021b).

Proponents emphasize that this fixity of our DNA sequence means that reverse causality from behavior or environmental exposures to the genome can be ruled out. Given this, genetic data can serve as exogenous measures of individual characteristics, which do not change over the life course, "facilitating the tracing of developmental paths" or as a "fixed point from which to observe child development" (Belsky & Israel, 2014; Harden et al., 2020). Scholars have argued that PGSs can be used as a "molecular tracer": "Just as a radiologist might administer a radioactive tracer to track the flow of blood within the body, researchers can use genetics as a molecular tracer to get a clearer image of how students progress through the twists and turns of the educational system" (Harden et al., 2020).

4.3 Gene–environment interplay

PGSs are also advertised as a more direct and powerful tool to explore how gene–environment interplay influences social outcomes. Broadly, gene–environment interplay with PGSs can be demarcated into three broad types: (1) PGS–environment interactions (e.g., does gender suppress "genetic potential" for educational attainment; Herd et al., 2019), (2) PGS–environment combinatory effects (e.g., how do "nature" and "nurture" combine to shape children's resemblance to their parents in human capital accumulations over time; Harden & Koellinger, 2020), and (3) PGS-through-environment pathways (e.g., through what social–psychological mechanisms does the education PGS increase educational attainment; Bolyard & Savelyev, 2020).

Proponents have argued that PGSs can reinvigorate the study of gene–environment interactions (G × E) with "robust measures of genotype," in contrast to the limited candidate G × E approach (Harden & Koellinger, 2020; Martschenko et al., 2019). "By applying the prism of GxE models, it is hoped that the white light of average effects will be refracted into a rainbow of genetically mediated responses that are made clear to the scholar interested in describing human behavior" (Conley, 2016, p. 293). In addition, PGSs may also be gainfully employed in the service of understanding heterogeneous responses to social interventions, in the form of a PGS × intervention (Harden & Koellinger, 2020).

4.4 Risk stratification and/or early identification

Although most scholars agree that PGS-based personalized programs or policies are not realistic because of poor individual

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

prediction, PGSs are still advertised as having potential use in risk stratification, particularly for those in the upper and lower deciles of PGSs. On this view, PGSs could be used to identify "at-risk" individuals before problems manifest or become severe through the implementation of an early genetic screening system (Martschenko et al., 2019). Such genetic screening is argued to provide an inexpensive way to more expansively identify those at high genetic risk of problems, such as lower educational attainment or physical inactivity, and intervene in advance with, for example, extra support or placement into a different learning environment (Harden & Koellinger, 2020; Martschenko et al., 2019). Similarly, PGSs could be used to identify "high potential" individuals, who could also be targeted with different learning environments.

In addition to risk stratification, proponents argue that enhanced understanding of the distribution of genetic risks could be used to study the effects of social institutions and programs. For example, in educational systems, studying the distribution of genetic risks "across schools could be used to study inequities in the current ways that the educational system under- and overdiagnoses students… thereby identifying differential diagnoses and treatment across groups" using PGSs as "indicators with some degree of objectivity" (Martschenko et al., 2019).

4.5 Changing worldviews and approaches to social inequalities

Finally, some proponents claim that incorporating genetics into social science will change the way that social scientists think about the world. In the words of Harden and Koellinger (2020, p. 567):

> Ultimately, the greatest impact from integrating genetics into the social sciences will probably not come from simply applying new tools to old questions, but from changing how people think about the world around them, allowing them to ask new questions and to pursue new answers that would not have been feasible before. For example, the realization that success in life is partly the result of a genetic lottery raises new questions not only about underlying mechanisms, but also about fairness and what a desirable distribution of wealth in a society should look like.

On this view, GWASs and PGSs reveal the hitherto unrecognized fact that "success in life" is partly shaped by our genetic inheritances. In general, these scholars maintain that incorporating genetics into social science will stimulate new ways of thinking about and investigating our differences and inequalities, which may inform social policies to ameliorate inequalities.

4.6 Summary

Proponents tout several benefits from incorporating PGSs into social science to enhance social science research. In the next section, I scrutinize the science of sociogenomics, highlighting limitations, which I argue, undermine the utility of PGSs into social science. Most of these limitations are acknowledged by sociogenomicists; yet the full implications of these challenges are invariably unheeded in practical applications.

5. Limitations of PGSs that undermine their utility for social science

As is well known, a person's social traits emerge from a complex interplay of environmental and genetic influences over their lifetime. As I have discussed, the goal of GWASs is to identify variant substitution effects as causal genetic effects, and the primary raison d'être of PGSs is to index genetic influences on (differences in) phenotypes. Proponents hype the value of PGSs for "unbraiding" and "disentangling" the effects of genetics and environments in shaping individual differences in complex social outcomes. Naturally, this only works if (a) genetic and environmental influences on traits can be differentiated, and, if so, (b) PGSs are relatively accurate and unbiased estimates of genetic influences (Barton, Hermisson, & Nordborg, 2019). Unfortunately, for a variety of biological, statistical, and developmental reasons, GWASs cannot disentangle "genetic" from "environmental" influences, such that PGSs do not index genetic influences on complex traits (Haworth et al., 2019; Morris et al., 2020a). In particular, dynamic population phenomena induce confounding between genotypes and complex social outcomes at multiple levels, *inter alia*: family, neighborhood, peer group, region, culture, nation, and historical time (Barton et al., 2019; Lawson et al., 2020). I discuss four primary limitations of PGSs that vitiate their utility for social science as measures of "genetic influences on" or "genetic propensities for" complex social traits: relatedness confounding, downward causation, limited coverage of genetic influences, and context-specificity.

5.1 Relatedness confounding of PGSs

The most widespread and widely recognized form of environmental confounding is because of (genetic) relatedness and passive gene–environment correlations. Basically, people who are more genetically similar (i.e., more closely related, even distantly) also tend to develop in more similar sociocultural, political, and physical environments, which influence most complex social traits. Thus, genotype and environments are correlated for non-causal reasons. Generally, relatedness confounding is demarcated into population genetic structure and familial confounding. Both are known issues in GWASs/PGSs and steps are taken to mitigate this confounding. However, evidence is mounting that these corrections are insufficient, such that inflated or spurious genetic associations persist (e.g., Barton et al., 2019; Berg et al., 2019; Haworth et al., 2019; Morris et al., 2020a; Mostafavi et al., 2020).

5.1.1 Population (sub)structure and phenotype stratification

> With respect to confounding by population structure, the key qualitative difference is between controlling the environment experimentally, and not doing so. Once we leave an experimental setting, we are effectively skating on thin ice, and whether the ice will hold depends on how far out we skate. (Barton et al., 2019, p. 3)

Population (genetic) (sub)structure refers to patterns of genetic variation within populations because of non-random mating. Population structure arises because of complex demographic histories (separation, migration, admixture), which result in mostly random allele frequency differences between population subgroups (Cardon & Palmer, 2003; Lawson et al., 2020). When these coarse population genetic subgroups (shaped by geographic region, race/ethnicity, social class, religion) are differentially exposed to trait-associated sociocultural and physical environmental factors – as they often are – alleles associated with subgroup membership are also associated with trait differences, producing spurious or inflated genetic effect size estimates, known as *phenotype stratification* (Browning & Browning, 2011; Cardon & Palmer, 2003; Morris et al., 2020a).

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

The classic example used to illustrate phenotype stratification is a genetic association study of chopstick-eating skills (Hamer, 2000; Lander & Schork, 1994). If we were to conduct a GWAS of using chopsticks in a sample of diverse ancestry, we would no doubt find significant associations. Although there may be some genetic variants affecting our ability to handle chopsticks (e.g., finger dexterity), most genetic associations would be because of cultural differences, namely random variants that differed in frequency between East Asia and the rest of the world and had nothing to do with "genetic propensity" for chopstick use skills. In practical applications, phenotype stratification is most plainly manifest with the geographic patterning of PGSs, which reflects sociocultural and physical environmental influences (Abdellaoui, Verweij, & Nivard, 2022; Haworth et al., 2019; Lawson et al., 2020).

The minimal approach to mitigate phenotype stratification is the examination of an ostensibly homogenous ancestral group. However, population substructure exists within these groups, including populations from a single location, such as "white Europeans" within the United Kingdom, Finland, the Netherlands, and Western France (e.g., Bycroft et al., 2019; Byrne, van Rheenen, van den Berg, Veldink, & McLaughlin, 2020; Haworth et al., 2019; Karakachoff et al., 2015; Kerminen et al., 2017; Leslie et al., 2015). Such finer-scale genetic population structure (known as local or regional population structure) is a function of non-random mating shaped by sociopolitical forces, cultural factors, and different physical environments all of which foster assortative mating (Morris et al., 2020a; Richardson & Jones, 2019; Zaidi & Mathieson, 2020). Consequently, pervasive, albeit often subtle, allele frequency differences between subgroups experiencing many different physical and social environments exist and can be picked up by GWASs as genetic causes, even if functionally unrelated to trait variation. For these reasons, in the presence of population structure, GWAS SNP associations may just be proxies for (or inflated by) an environmental variable that has not been properly corrected (Browning & Browning, 2011; Cardon & Palmer, 2003; Novembre & Barton, 2018).

Several sophisticated statistical methods have been introduced to mitigate or adjust for population structure-confounding, including genomic control (Devlin & Roeder, 1999), genetic principal components (PCs) (Price et al., 2006), linear-mixed models (LMM) (Kang et al., 2010), and LD score regression (LDSC) (Bulik-Sullivan et al., 2015). Although these methods appear to reduce population stratification, evidence from a variety of studies using whole-genome sequence data, simulations, and tests of non-genetic traits (like latitude/longitude of birth, birth order) evince that these methods do not adequately correct for population structure, and this is especially true for complex social traits of interest to sociogenomicists (e.g., Berg et al., 2019; Dandine-Roulland et al., 2016; Mostafavi et al., 2020; Sohail et al., 2019; Zaidi & Mathieson, 2020).

For example, in a recent study, Abdellaoui et al. (2022) demonstrate that controlling for geographic region decreases heritability signals for socioeconomic status (SES)-related traits, especially educational attainment and income, as socioeconomic differences between geographic regions induce gene–environment correlations that are picked up in GWASs and inflate PGSs (see also Leslie et al., 2015; Mostafavi et al., 2020; Sohail et al., 2019). In another study using simulations, Zaidi and Mathieson (2020) show that recent (within the past 100 generations or ~2,500 years) genetic structure with sharp effects pose a particular problem for GWASs/PGSs given the tag SNP methodology. As they explain, recent population structure with sharp local effects, as may result from cultural, language, and/or physical boundaries patterning mating, can only be adequately corrected with rare variants, which are not measured in these studies.[9]

In sum, the evidence is clear that phenotype stratification persists despite sophisticated methods to mitigate such confounding – most obviously in the form of geographic patterning of PGSs (Abdellaoui et al., 2022; Byrne et al., 2020; Haworth et al., 2019) – and its effects (inflating PGSs) appear to be particularly acute for complex behavioral traits related to socioeconomic status (Abdellaoui et al., 2022; Lawson et al., 2020). Crucially, these biases are exacerbated under the very modeling conditions most often used for social science outcomes – when multiple studies are meta-analyzed and millions of SNPs are aggregated in PGSs. Even subtle population stratification can cumulatively generate substantial biases when millions of SNPs are aggregated, especially when less stringent *p*-values are employed (as is typical) (Barton et al., 2019; Berg et al., 2019; Mathieson & Mcvean, 2012). In short, PGSs for complex social traits capture some non-trivial amount of social environmental effects because of uncorrected population substructure (Abdellaoui et al., 2022; Curtis, 2018; Lawson et al., 2020).

5.1.2 Familial confounding[10]

Biological parents not only pass on half of their genome to their children but also their environments, including social status, culture, worldviews, values, habits, and the like (Shen & Feldman, 2020). Therefore, the association between parental and offspring genotypes is often confounded by the association of genotypes with rearing environments, effects which may be amplified over generations via social mechanisms (as "dynastic effects"; Brumpton et al., 2020). Such gene–environment correlations inflate estimates of genetic influences, especially for complex social traits where the transmission of social advantages (e.g., status and wealth) and associated familial practices are significant (e.g., Kong et al., 2018; Morris et al., 2020a).

Several innovative sociogenomics studies have illuminated the extent of familial confounding in PGSs. These studies suggest that roughly half of the effect of the education PGS is because of familial confounding. For example, Kong et al. (2018) found that controlling for an education PGS created from parents' non-transmitted alleles (i.e., the other half of alleles not passed down) reduced the variance explained by the offspring education PGSs by roughly half. If child PGS captures causal genetic effects, then controlling for non-transmitted parental alleles would not substantially reduce the effect of the child PGS on their education. In contrast, Kong et al.'s results suggested significant inflation of ostensibly genetic effects by familial confounding. In another study, Cheesman et al. (2020) compared the predictive effects of an education PGS on years of education in adopted and non-adopted youth. They observed that the PGS was twice as predictive of years of education in non-adopted versus adopted individuals ($R^2 = 0.074$ vs. 0.037), as would be expected if the education PGS captures familial effects. Similarly, Belsky et al. (2018) observed that controlling for parental education reduced the effect of the education PGS on years of education by about half, which "suggests environmental confounding of polygenic score associations with educational attainment" (p. E7277).

As with population structure, practitioners are aware of the issues with familial confounding and have employed statistical techniques to attempt to mitigate this confounding (see, e.g., Trejo & Domingue, 2019; Wu et al., 2021; Young et al., 2018). The most rigorous approach to reduce familial and population

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

structure confounding is a within-family or sibling-difference design. These studies examine how differences between siblings in their genotypes (in GWAS or PGS prediction) explain sibling differences in phenotypes, net of their shared rearing environments using family fixed effects (Belsky et al., 2018; Laird & Lange, 2006). For illustration, Lee et al. (2018) used a sibling-difference study to test the robustness of their (conventionally) unrelated sample education GWAS findings using a sample of ~22,000 sibling pairs. Given differences in statistical power, Lee et al. (2018) examined sign concordances of the GWAS coefficients (i.e., whether the effect direction of the risk alleles matched +/+) rather than their significance or effect sizes across the studies at three different *p*-value thresholds. By chance, of course, we would expect 50% of the signs to match. Their results showed that for the less stringent *p*-value threshold ($p < 5 \times 10^{-3}$), sign concordances between the discovery GWAS and sibling-difference GWAS were only slightly better than chance at ~56.5%, which improved at more stringent *p*-value thresholds to ~60% at $p < 5 \times 10^{-5}$ and ~65% at $p < 5 \times 10^{-8}$.[11] [Aside: Although expecting perfect sign concordance is unrealistic, a sign concordance of <57% at a *p*-value threshold that was *more* stringent than the one employed to create the widely used education PGS does not, in my view, demonstrate robustness or constitute replicated findings.] Lee et al. (2018) reported that the within-family effect sizes were, on average, 40% smaller than that from the unrelated GWASs. The just-published updated education GWAS did not present a within-family GWAS replication; however, their within-family PGS analyses indicated that only 30.9% of the PGS effect was a "direct effect" (Okbay et al., 2022; see also Morris et al., 2020a).

Not unexpectedly, sibling-difference studies of non-social (more proximally biological) traits, like height and C-reactive protein, report only minor evidence of familial confounding and slightly reduced effect sizes, whereas sib-studies of social outcomes, like educational attainment and smoking behavior, invariably report appreciably smaller effect size estimates, given the significance of sociocultural forces on these traits (Howe et al., 2022; Lee et al., 2018; Mostafavi et al., 2020). Importantly, this confounding is not simply a minor issue affecting the precise effect size but evidence suggests that this confounding substantively alters sociogenomics findings. For example, Howe et al. (2022) demonstrated that strong genetic correlations between education and height, weight, and C-reactive protein from population genetic studies become "negligible" in sibling-difference analyses.

Given the persistence of genetic relatedness confounding in GWASs and PGSs even with sophisticated methodological "corrections," research employing PGSs as indicators of genetic influence should, at a minimum (a) control for relevant social environments that are associated with genotype, including geographic location (Abdellaoui et al., 2022), or, preferably, (b) use sibling-study adjusted PGSs through a two-stage model to reduce (if not completely eliminate[12]) relatedness confounding. In the two-stage model, SNP *p*-values are estimated using a large unrelated GWAS, but the effect sizes are adjusted (downward) using the coefficients from a sibling-difference study (Choi et al., 2020; Zaidi & Mathieson, 2020). Unfortunately, neither is common practice. Estimates used to create the education PGS, now widely available for use in social science datasets, were not adjusted based on the sibling study reduced effects sizes or the sign mismatch in the replication mentioned above. Creditably, the authors (Lee et al., 2018) recognized the persistence of confounding, writing:

> [o]ur within-family analyses suggest that GWAS estimates may overstate the causal effect sizes: if educational attainment-increasing genotypes are associated with parental educational attainment-increasing genotypes, which are in turn associated with rearing environments that promote educational attainment, then failure to control for rearing environment will bias GWAS estimates…. *Without controls for this bias, it is therefore inappropriate to interpret the polygenic score for educational attainment as a measure of genetic endowment* (p. 1116, emphasis added).

Despite this clear caution about using PGSs as genetic potential without controls for confounding, subsequent education PGS studies did not heed these cautions and failed to control for rearing environments while examining PGSs as "genetic propensity" (e.g., Harden et al., 2020; Herd et al., 2019; Wedow et al., 2018).

Notably, even PGSs created from within-family GWASs are not immune to environmental confounding for two key reasons. One has to do with the uniqueness of within-family designs. Because of subtle micro-stratification and complex social–psychological dynamics within families, the extent to which the causes of sibling differences for complex social traits are the same as the causes of general population differences is questionable. Research suggests sibling differences may be amplified or distorted as siblings attempt to create their own niches or fill unique roles in their families (e.g., "the smart one," "the athlete," "the funny one," "the troublemaker," "the pretty one") (see, e.g., Healey & Ellis, 2007; Sulloway, 2001) in part through "sibling contrast effects" (Carey, 1986). For other traits and behaviors, differences may be minimized as families tend to socialize children in similar ways and siblings imitate one another. These interactional dynamics influence child identities, expectations, motivations, personality, and developmental outcomes and thus undermine the generalizability of sibling-difference studies.[13]

In addition, genetic associations and PGSs from sib-studies are confounded by broader sociocultural influences. This is because the counterfactual model that underlies genetic association studies does not distinguish between authentic (upward) genetic causes (i.e., from genetic differences to trait differences through biological mechanisms) and artificial downward (social) causation. Both are identified as causes in GWAS's counterfactual variant substation effects approach.

5.2 Downward causation and artificial genetic signals

Downward causation – defined as sociocultural forces that sort and select individuals based on genetically influenced traits, such as skin pigmentation and height, into different environments and exposures that influence social outcomes – creates what I call *artificial genetic associations*, which are environmental influences masquerading as genetic influences in GWASs. Although the fact that sociocultural environments shape and filter genetic influences is understood by most, less well understood is the extent to which the causal effects of social structural and cultural forces acting on genetically influenced differences are identified as genetic influences in GWASs and PGSs.[14]

Jencks' (1972) now classic thought experiment on discrimination by hair color can be used to illustrate downward causation creating artificial genetic associations. Jencks asks us to imagine a system where red-haired children are barred from school. In such a system, genetic variants linked to red hair would be identified by GWASs as genetic causes of educational attainment. However, neither an individuals' red hair, nor the genetic variants contributing to red hair, are appropriately conceived as causes of

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

differences in educational attainment in this hypothetical case, in my view and that of others (Kaplan & Turkheimer, 2021), but see Harden (2021a). The "difference that makes a difference" is not red hair but the social-institutional policies excluding people with red hair, which is why a change in the rules would (over time, I presume) make hair color unrelated to educational attainment (and remove any red-hair genetic associations with education). Although explicit discriminatory exclusionary policies like this one are largely a thing of the past in most developed nations, both ongoing discrimination and the legacy of past discrimination (through intergenerational transmissions of wealth, status, social capital, etc.) continue to influence individual development and trait differences. More broadly, our environments and institutions, educational and otherwise, continue to differentially treat individuals based on a variety of genetically influenced individual traits such as height, body weight, personality, attractiveness, and skin tone into different environments and exposures and thus opportunities, achievements, and developmental outcomes (e.g., Monk, Esposito, & Lee, 2021; Simons, Burt, Barr, Lei, & Stewart, 2014).

GWASs and PGSs capture artificial genetic signals, and these artificial effects are likely to be pervasive given the extent to which we respond to phenotypic cues in our interactions with others in a manner that is unavoidably socioculturally mediated. Although casting such socioculturally driven genetic associations as genetic propensity or even "indirect genetic effects" is misguided, even more concerning is the subsequent framing of such correlations as innate individual propensities (individual "genetic fortune" or "misfortune"). Because of downward causation, genetic associations for many complex social behaviors are unavoidably environmentally confounded and are not appropriately conceived as genetic causes of outcomes.

5.3 Limited coverage of genetic variation

To serve as a control for genetic influences, in addition to not being substantially environmentally confounded, PGSs need to capture genetic influences relatively accurately and comprehensively. They do not.

5.3.1 Low resolution

GWASs and PGSs capture genetic variation at low resolution. As noted, SNPs rarely have functional effects and usually tag large regions of common variation, which may contain numerous causal variants including large effect extremely rare variants (McClellan & King, 2010).[15] The causal variant(s) in the tagged region may often be multiple and rare, and such that only a paucity of individuals with the risk allele (tag SNP) will carry the actual causal variant. Thus, tag SNPs – even if they reflect causal genetic influences – are very imprecise proxies for a causal variant that may only exist on that haplotype for a small minority of individuals.[16] The tag SNP methodology, which excludes rarer and likely functional SNVs, indels, and SVs make GWASs possible, but it also makes PGSs incomprehensive measures of genetic risk (Backman et al., 2021).

PGSs also ignore the X chromosome (given that females have two and one is usually inactivated in a cell), and both GWASs and PGSs invariably ignore the Y chromosome. Mitochondrial DNA is also neglected.

5.3.2 Genetic additivity and interactionism

Finally, GWASs and PGSs usually estimate additive genetic influences. However, because of pervasive gene–gene interactions and interactions between non-coding RNA genes and coding genes, focusing on additive effects from tag SNPs is necessarily misleading (as oversimplified) about the true nature of genetic influences (Belsky & Israel, 2014; Zuk, Hechter, Sunyaev, & Lander, 2012). Almost everything that happens even at the cellular level is because of the combined influences of different molecular mechanisms, such as different proteins and functional RNA molecules. Given that, the idea that genotypes can just be summed together to arrive at a measure of genetic liability seems naïve.

To be sure, evidence for a substantial role of interactionism is lacking; however, the current evidence is primarily based on low-resolution tag SNP methodologies. That low-resolution methods have not yet substantiated the importance of gene–gene interactions, does not suggest they are not biologically important.

In sum, for a variety of methodological reasons, PGSs do not control for genetic heterogeneity. The final limitation of PGSs I consider relates to the neglect of developmental interactionism. As I discuss next, the well-known context-specificity of genetic influences (Feldman & Lewontin, 1975) impedes some of the intended uses of PGSs.

5.4 Context and population specificity

That heritability studies are context- and population-specific – a point made clearly and forcefully by Lewontin (1974) nearly 50 years ago – is now widely appreciated after considerable scholarly effort and some costly misrepresentations (Jensen, 1967). However, that GWASs and PGSs are similarly context- and population-specific is not as widely appreciated in theory or practice (but see Kaplan & Turkheimer, 2021). It should be. This is particularly true for non-biological social behaviors and achievements like educational attainment or same-sex sex, which involve somewhat arbitrary institutional structures (e.g., financial resources and opportunities) as well as cultural norms.[17] For reasons expounded upon below, such genetic associations should not be understood as timeless, context-independent genetic influences. That is, even if we could disentangle the influence of genes from environments for these outcomes, these associations reflect developmental gene–environment interactions under current social arrangements in each context, not what could be in different circumstances (historical periods, social position, cultural context, etc.).

This well-known context- and population-specificity exists for two general reasons. The first is biological: Genes always interact with environments across all levels of development in their effects on complex traits. The second is sociocultural: The individual characteristics influencing traits or achievements, and thus the genetic contributors thereto, vary across historical time, society, and even across structural location. For illustration, the genetically influenced individual traits facilitating educational attainment for a woman in Saudi Arabia in 2000 versus a woman in 1870s in United States, in 2010 in India, in 2002 in Nigeria, in 1950 in Thailand, or in 2021 in United States are likely to be distinct in non-trivial ways. Although a woman going to college in the United States in 2020 would be conforming, a woman going to college in 1870s in United States would be statistically deviant. Because educational attainment reflects numerous genetically influenced traits, filtered by context and relative condition, the idea of a context-invariant "genetic propensity to" complex social outcomes like educational attainment, like crime, smoking, or same-sex sex, is misguided (Burt, 2023).

Moreover, the search for a "winning" genetic endowment that can be measured on a unidimensional scale representing propensity for social success is also misguided, in my view (e.g., Belsky et al.,

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

2016). This is because our DNA is part of an interactional developmental system that responds to context- and condition-dependent stimuli (Burt, 2018; Ellis et al., 2012). Genetic differences influencing complex traits, like traits themselves, are not amenable to facile "good" or "bad," "winning" or "losing" ratings but rather more like "it depends," on a host of other factors (e.g., other genetic differences, other traits, historical context, social class, etc.). To use an oversimplified example, while being confident, independent, and talkative may enhance educational attainment and occupational success for an upper-middle class white male, those same traits among a minority youth from a disadvantaged background could very well impede educational attainment. Of course, confidence and independence emerge from a host of influences, but the point of this example is to reveal the oversimplified (theoretically and empirically unwarranted) model underlying an additive genetic index representing a context-independent propensity for complex social behaviors like educational attainment.

The problems with a unidimensional genetic propensity for complex biological traits are even more obvious for a phenotype of (having ever had) same-sex sex. As with people who attain higher levels of educational attainment, people who have ever had same-sex sex display remarkable diversity. From "gold star" lesbians and bisexual women to "femme" women who have same-sex sex only to please their male partners, the search for an additive, context-independent underlying continuum of genetic propensity for "having ever had same-sex sex" is empirically and theoretically unwarranted. Not only is there expansive heterogeneity within these groups, but also same-sex sex, like other social behaviors such as doing ballet, trying ecstasy (MDMA), and playing golf, is not simply the outer manifestation of some inner potentiality. Different sociocultural constraints and opportunities shape the behavioral manifestation of various traits and propensities, however genetic, which are then further altered by social responses in developmental feedback loops (including labeling and self-identification). Of course, we can impose a unidimensional propensity measure – a PGS or otherwise – for such heterogeneous and socially contingent behaviors by estimating the probability of the binary measure of having ever done so. But creating such a continuum statistically does not mean such a propensity exists biologically.

Thus, for yet another reason, PGSs cannot be thought of as "genetic potential," inasmuch as genetic influences are not static charges where PGS effects sizes can be facilely compared across contexts or conditions. Traits that facilitate educational attainment, and any genetic contributions thereto, are dependent on sociocultural influences. For example, if physical education classes were equally emphasized with non-PE courses and graded not by effort but also by achievement, academic attainment may look noticeably different.

This context-specificity has implications for some prominent applications of PGSs. Following prior behavioral genetics work that examined how heritability estimates varied across contexts or conditions, several recent studies have used PGSs to explore how "genetic influences" are moderated by (often "constrained" or "suppressed" in) different contexts or for different social groups (Harden et al., 2020; Trejo et al., 2018; Wedow et al., 2018). For example, Herd et al. (2019) examined whether "the influence of genetics on educational attainment has changed across cohorts" and "whether this influence varies by gender" by comparing the effect sizes of the education PGS on educational attainment across cohorts (defined by historical time) and by sex. Their focal hypothesis was that among older cohorts, social structures of gender suppressed the "genetic potential for educational attainment" among women but not men, manifest in weaker education PGS prediction among women in older cohorts. To be sure, the Herd et al. study was explicitly sensitive to context, recognizing how genetic effects are "filtered, altered, and shaped by broader complex environments" (p. 1071). Even so, this approach remains insufficiently context-situated and oversimplified. This is because the study rests on the idea that PGSs capture a historically invariant genetic potential for educational attainment, such that weaker PGS prediction can be interpreted as lesser genetic influence and thus suppressed potential. However, for reasons mentioned above, as contexts and opportunities change, so too do the characteristics influencing achievements and social behaviors, and thus their genetic influences. A weaker PGS across contexts may just mean different traits matter (and would be expected in this example for statistical reasons given the lower mean and variance of educational attainment in the earlier cohorts compared to the latter ones). For all these reasons, interpreting effect size differences in PGSs as indicating that "genetic influences matter less" for social traits in different contexts or as evidence that "potential is suppressed" is unsound.

Upon deeper reflection, the extent to which research into how contexts suppress or constrain "genetic potential" (via reductions in PGS effect sizes) advances knowledge is unclear. Leaving aside my objection to the notion of a context-independent genetic potential for social traits, in general, and PGSs as an indicator of such potential, in particular, what, specifically, is the value of assessing whether "genetic potential" is suppressed by these social arrangements? Until well into the twentieth century, the potential for educational attainment for women in the United States was, of course, constrained by structures of gender that limited them to family roles in the household. We already know women's potential was suppressed, in these instances. What would it mean to say that potential was suppressed but not genetic potential? Is the null hypothesis that only "non-genetic potential" was suppressed (and what would that even mean)? Phrased alternatively, given that potential emerges from developmental systems shaped by interacting genetic and environmental forces, is there any argument that can be made that discriminatory arrangements or disadvantages constrain achievement but do not affect genetic potential? How would that work?

6. Questioning substantive value added

Even if the problems with environmental confounding could be solved, the justification for incorporating PGSs into social science is lacking. The scientific warrant to include PGSs to reveal well-established social patterns more precisely or rigorously is, in my view, wanting. Given that we have robust evidence that higher education is associated with higher income, fewer children, and better health, what is the value of demonstrating that an education PGS is associated with fewer children born, household wealth, or health? How could it not be? A recent study with an education PGS investigated whether "parental genetics for educational attainment" are associated with better (i.e., warm, stimulating) parenting, thereby partially explaining the association between parents' education PGS and youth educational attainment (Wertz et al., 2019). Armstrong-Carter et al. (2020) highlighted this study as illustrating how "genes can be used as a lens for the study of social processes through which parents influence their children." Do we need GWASs, PGSs, and studies of "genetic nurture" to demonstrate that supportive, stimulating parenting is

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

associated with child educational attainment and that higher educated – disproportionately well-off – parents are more likely to engage in such parenting? Or that "children who experience childhood disadvantage are not able to fully realize their educational potential" (Ronda et al., 2020). Or that "that genetic endowments linked to educational attainment strongly and robustly predict wealth at retirement" (Barth, Papageorge, & Thom, 2020). I think not.

Harden et al. (2020) touted the potential of PGSs as "molecular tracers" for social achievements, like educational attainment, that can "measure flows of students through the STEM pipeline and assess how these flows differ across schools" analogous to how "a radiologist might administer a radioactive tracer to track the flow of blood within the body." However, the reason that radiologists use molecular tracers to trace internal functions is because they cannot observe such internal bodily processes. Unlike the radiologist tracking unobservable internal bodily processes like blood flow, we can observe and measure different student aptitudes, skills, and background factors and assess how these affect student progressions through educational systems. Given that opportunities exist for measuring background factors and proximal behaviors and that we already have a glut of assessments (e.g., grades, cognitive testing), the need for and utility of such a tracer – which those scholars admit is not a useful individual predictor – is surely questionable (Morris et al., 2020b).

In addition to meager benefits, such research has several potential costs. The use of PGSs as molecular tracers is rooted in the misguided idea that PGSs reflect individual propensity – that is, that the potential for educational success resides in our genome. Indeed, the authors argue that "[t]his approach offers a way of diagnosing the extent to which students who have *high genetic propensities for success in education* leak out of the STEM pipeline by failing to advance in their mathematics training" (Harden et al., 2020; emphasis added). Not only are PGSs flawed as measures of "high genetic potential" but the concern with the "high genetic potential" students "leaking out of the STEM pipeline" seems unjustified given paltry PGS individual prediction and the fact that potential for complex social achievements like years of education cannot be reduced to genotype (which the authors acknowledge). The paper evidences a heightened concern over the "high genetic potential" students leaking out over their "lesser potential" (lower PGS) counterparts, but this concern is never explained. Even more concerningly, this focus on the "high genetic propensity" seems to reflect the privileging of the purportedly "genetically gifted" in a manner that will increase rather than decrease inequalities.

To be sure, Harden et al. (2020) highlight the potential of the education PGS as a molecular tracer to inform school performance evaluations with the explicit aim of ameliorating inequality. However, such applications of school-level "genetic potential" performance assessment would, given existing social arrangements and environmental confounding, identify schools with a much higher proportion of lower income students from less-educated families as having lower *genetic* potential. Using PGSs as potentials, schools with such lower performing students would thus not be identified as "underperforming" because their students just "lost" in the "genetic lottery" (and we cannot expect much from them on this view). Although this is clearly not the intention of the authors, using PGSs as tracers necessarily rests on the idea of PGSs as indicating genetic potential for educational success – and, as noted, the authors use such terminology.[18] Casting PGSs as "potential" risks reifying genetic differences among groups with different social behaviors and attainments shaped by prior and existing unequal arrangements as "genetic potential" and then excusing future patterns as inevitable because of genetic propensities, even for traits that are substantially driven by social inequalities and malleability.

These studies are in no way unique among sociogenomics studies but instead reflect the implicit "because we can" rationale of much sociogenomics research, often evidenced by the wholly uncompelling justification for some studies. Take the GWAS of "having ever had same-sex sex." Ganna et al. (2019) explain the value of their study as follows: "With respect to genetic influences [on same-sex sex], several questions arise. First, what genes are involved and what biological processes do they affect? … Identification of robustly associated variants could enable exploration of the biological pathways and processes involved in development of same-sex sexual behavior" (p. 1). Leaving aside the implicit assumption of a molecular pathology underlying 'non-heterosexuality' indicated by "having ever had same-sex sex," as we have discussed, GWASs are not at all well suited for identifying genes, underlying causal variants, or tracing biological pathways for complex traits. In short, that scholars can conduct a study, does not mean that they should (i.e., that doing so advances science).[19]

From a broader perspective, sociogenomics' ambiguous contributions to knowledge are because of a prevailing deficit of theory, especially as relates to causal theories about developmental processes, which permits a rather shallow approach to the meaning of genetics plus social questions. To be sure, that social science genetics has a deficit of theory is not a novel criticism (e.g., Boardman & Fletcher, 2021; Burt, 2022; Panofsky, 2014), but attention to this neglect of theory and the manner in which this neglect hampers knowledge advancement is scarce. In my view, excitement over our ability to conduct analyses with incredibly advanced statistical and genetic tools appears to overshadow limitations and a sober evaluation of limitations. All too often, the contemporary enthusiasm around applying new genomics tools to social science adds a sheen that glosses over the meager practical and scientific contributions of this work, beyond simply showing that PGSs are statistically significant or have some non-trivial R^2.[20] At this point, no serious scientist can suggest that genetic differences do not influence – in some complex, context-dependent way – developmental differences. Simply demonstrating that yet again with sophisticated, albeit biased, methods does not advance understanding (see also Turkheimer, 2016).

Finally, as noted, scholars point to PGSs as a control to "get genetics out of the way" to reveal aspects of our environment; however, I have yet to see any sociogenomics findings that change our understanding of environmental influences or suggest different policy or programmatic approaches. Given the limitations mentioned above, I am unable to conceive of any research findings at the present state of the science, which would support such changes in theory or practice. That is, even if the inclusion of PGSs markedly altered an environmental estimate, because PGSs are significantly environmentally confounded, we cannot say that controlling for "genetics" is the cause of such changes. What is more, we cannot say that environments matter "net of genetics" because PGSs only capture a fraction of the ostensible heritability of social outcomes. What, then, can or should we do? Below, I outline suggestions for sociogenomics at the current state of the science.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

7. Suggestions

An abundance of genetic data is available for incorporation into social science with increasingly advanced computational methods and enhanced rigor in approach, relative to earlier eras. Given the limitations I have discussed along with my arguments about limited contributions, how should PGSs be used in social science, in my view? My answer is quite possibly unsatisfying: Sparingly and cautiously with caveats placed front and center. Enthusiasm about the opportunities genetics offers behavioral science should be tempered with a more realistic appraisal of current challenges and uncertainties. After all, we have been here – with excitement around genetics, limitations in methodology, and substantial unknown biology – before, quite recently, with the candidate gene era of a few years ago (see Charney, 2022).

Scholars should be more skeptical of the value added of PGSs to social science, and I have several suggestions to this end. First, when considering incorporating PGSs, behavioral scientists should first ask whether the outcome is a sufficiently tightly biologically regulated phenotype amenable to molecular genetic analyses. If so, scholars should explicitly specify how incorporating genetics advances science with a sufficiently high bar, one which acknowledges potential risks and benefits and recognizes that it is already well established that our genetic differences do matter in a complex, context-sensitive way (Turkheimer, 2016). Simply "showcasing the power of genetics" by revealing that PGSs are correlated with some outcome does not advance knowledge. Additionally, sociogenomics research should include controls for social variables associated with complex traits. At present, all too often easily measured and relevant social science predictors are not included in research "showcasing the power of genetics." This is unsatisfactory.

Importantly, sociogenomics scholarship should eschew terminology that implies that genetic differences are driving behavioral differences given pervasive and unavoidable environmental confounding for all social outcomes. Framing PGSs as "genetic influences" should be avoided, and terminology like "association" or "correlation" should be employed instead. Likewise, I urge scholars to avoid "propensity" terminology or treating genetic endowment as a "lottery" in which there are winners and losers for complex social outcomes. Even if we could identify genetic influences on, for example, the type of intelligence that facilitates educational success and wealth, facilely equating genotypes associated with such capacities to "winning" at genetic inheritance or, conversely a lower education PGS as "an unfavorable genetic endowment" (e.g., Bolyard & Savelyev, 2020), is misguided. That is, of course, not to deny that people with greater wealth have better health and easier times dealing with stressors, on average; rather, it is to say that neither higher education nor greater wealth equals winning "the good life," whatever that is.

In sum, I urge sociogenomics to think about where the science is, not where it might be (avoid hype and promissory notes); to acknowledge what questions we can answer at the current state of knowledge and which ones we cannot; and, finally, to recognize that just because social scientists can incorporate PGSs into our models, does not mean that we should – that is, that doing so advances knowledge.

8. Summary and discussion

Here, I challenged proponents' claims about the scientific warrant to include PGSs in social science. After outlining proponents' arguments about the utility of PGSs for social science, I argued that these ostensible scientific and practical benefits rely on the misguided notion that PGSs represent "genetic influences" on complex social traits. Instead, I explain that PGSs are unavoidably environmentally confounded because of population stratification, familial confounding, and downward (socio-environmental) causation. Although methods exist to mitigate the former, especially within-family studies, artificial genetic association signals created by downward causation cannot be differentiated from authentic genetic signals with the counterfactual models employed. In addition, I explain why PGSs do not, in fact, accurately or comprehensively control for "genetic influences" on traits because of methodological limitations (e.g., the tag SNP methodology) and biological challenges (including the nature of genetic influences). Finally, I discussed the context-specificity of PGSs, which precludes their use as "genetic potential" in general, and comparisons across contexts and conditions as a means of assessing the suppression of "genetic influences," in particular. I explained that these models remain fundamentally and necessarily wedded to an overly simplistic and ultimately misleading (environmentally confounded and biologically implausible) reductionist genes-versus-environments approach.

In response to this critique, scholars may point to the fact that "PGSs just work." By that, they presumably mean that PGSs "predict" the outcomes they were created to predict, even differences within families, albeit weakly in a manner that is inappropriate for individual prediction. However, the potential of PGSs is not rooted in their statistical predictive ability, however meager or substantial, but in their capturing genetic (vs. environmental) influences on trait differences. Furthermore, for complex social traits like education, as Morris et al. (2020b) documented in their evaluation of practical utility, an education PGS "provided little information on [youth] future achievement over phenotypic data that is either available or easily obtainable by educators."

Others may respond by suggesting that I am holding sociogenomics methods to higher standards than standard social science methodologies.[21] To that charge I cannot plead "not guilty." Instead, I justify my scrutiny by pointing to the prior missteps in social science genetics, including the recent spectacular failure of the candidate gene era, the incautious hype, and the potential for misuse (see Dick et al., 2015; Yong, 2019). Moreover, proponents and critics alike have recognized that the scientific and social risks for the misinterpretation of PGSs are real and potentially significant, a situation exacerbated by the media tendency to ignore caveats and uncertainties and social scientists' lack of expertise in genetics (Barton et al., 2019; Richardson, 2017). These risks behoove us to approach the incorporation of genetics into social science with special caution and appropriate scientific skepticism.

Whether and to what extent incorporating genetics can benefit social science theory and research in a manner that may have practical implications remains to be seen. In my view, the payoffs for studying genetic influences on non-disease complex social traits and achievements for most applications are minimal. The potential costs of prematurely and misguidedly promoting PGSs as "genetic potential" are significant, and include, in addition to wasting finite resources searching for "genes for educational attainment," obscuring social–structural and physical environmental influences and promoting the individualization of social problems.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

9. Caveats and conclusion

My critique is intended to promote a dialogue between social and behavioral scientists about the scientific value of adding genetics to social science at the current state of knowledge. I hope this discussion eschews hype, straw man arguments, imputing motives, and ad hominem – all of which foster misunderstanding, polarization, even hostility. If we avoid such discussion-impairing tactics, which characterized some prior efforts to discuss genetics in social science, both science and society will be the better for it.

To avoid misunderstanding, I wish to clarify that my stance does not imply that the incorporation of genetics into social science necessarily involves racist motives and/or tacit support for eugenics; it quite clearly does not. Moreover, this critique is not motivated by a desire to censure scholars by imputing (bad) motives or to censor areas of study for ideological reasons or because of sociopolitical concerns. My aim is to draw attention to limitations of incorporating PGSs into social science and misinterpretations with the aim of promoting better science.

In the end, my argument is simply that the claims made by proponents about the benefits of PGSs and their utility as measures of "genetic influences" or "genetic propensity" are overstated and misguided. Because of these limitations, PGSs cannot be employed as measures of "genetic influences" as they are being used with increasingly regularity. GWASs and PGSs may be powerful tools for identifying genetic associations, but they are not the right tools for understanding complex social traits.

Acknowledgments. I am extraordinarily grateful to Kara Hannula for her valuable comments on multiple drafts of this manuscript. I am also very grateful to the helpful guidance from Barbara Finlay and the detailed comments and suggestions from the anonymous reviewers, which significantly improved the manuscript. The content is solely the responsibility of the author and does not represent the views of the National Institutes of Health or those who provided feedback.

Financial support. Support for this research was provided by K01 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (5K01HD094999). Partial support for this research came from a Eunice Kennedy Shriver National Institute of Child Health and Human Development research infrastructure grant, P2C HD042828, to the Center for Studies in Demography & Ecology at the University of Washington.

Competing interest. None.

Notes

1. Notably, my coverage is not exhaustive. I highlight key issues, drawing selectively on scholarship in these areas given finite space. I do not discuss, e.g., the issue of selectivity (non-generalizability) of samples that predominant in GWASs (e.g., UK Biobank and 23&Me samples) (see, e.g., Burt & Munafò, 2021; Fry et al., 2017); the lack of ancestral diversity in genomic data; or what one reviewer called "the crude conceptualisation of psycho-social traits implicit in GWAS/PGSs and of the measures used."
2. Or 4–5 nucleotide differences every 1,000 bp accounting for structural variants.
3. A relatively small number of GWASs (but none in sociogenomics) have analyzed common copy number variants (CNVs) (see, e.g., Bochukova et al., 2010; Willer et al., 2009).
4. E.g., in their recent UK Biobank study using whole-exome sequencing, Backman et al. (2021) noted: "Rare variant associations were enriched in loci from genome-wide association studies (GWAS), but most (91%) were independent of common variant signals."
5. Commonly used reference panels include the 1KG, HapMap Phase 2, and, more recently, the ancestrally diverse Trans-Omics for Precision Medicine (TOP Med) sample (Taliun et al., 2021). For better or worse, the reference panels differ across samples used in meta-analyses. One might think it wise to control for the reference population used for imputation in a meta-analysis; however, I have not seen this done in practice.
6. As noted elsewhere (Burt & Munafò, 2021), these various thresholds are somewhat arbitrary and vary across studies, increasing, as others have also noted, researcher degrees of freedom (Charney, 2022).
7. As with the use of SNP associations for GWAS follow-up, when constructing PGSs, LD between SNPs needs to be accounted for to avoid aggregating SNPs that tag the same region of variation (i.e., multiple counting). That said, not all studies correct for LD when creating PGSs (see, e.g., Wertz et al., 2018, 2019). The consequence is an inflated PGS because of counting multiple SNPs that tag the same effect.
8. Some more sophisticated models, like LDPred, do not use *p*-value thresholds but instead involve the selection of various priors (assumptions) about the number of causal SNPs. In practice, the prior is that "all SNPs are causal," which is curiously not defended anywhere to our knowledge. Moreover, the idea that all SNPs have causal effects is not consistent with available empirical evidence.
9. Recent population structure is driven by rare variants which have a more recent origin and therefore are less likely to be shared among population subgroups (Fu et al., 2013; O'Connor et al., 2015). As such, recent structure (with sharper effects) cannot be captured by or corrected with common SNPs used in GWASs (Zaidi & Mathieson, 2020).
10. Familial confounding is sometimes called "indirect genetic effects" or "genetic nurture"; however, I eschew these terms because these imply a causal effect of parents' genotypes on child phenotypes through nurture, which has not been demonstrated. Familial confounding also includes so-called "dynastic effects" as (dis)advantages passed down to children (Abdellaoui et al., 2022).
11. These findings provide further evidence that the "all SNP"/no *p*-value threshold PGSs employed in most studies capture more bias than PGSs with *p*-value thresholds (Barton et al., 2019; Berg et al., 2019; Sohail et al., 2019).
12. Importantly, although sibling difference PGS studies significantly reduce environmental confounding, they do not eliminate it; as Zaidi and Mathieson explain, although estimates are unbiased, stratification in the PGSs persists because the frequency of the SNPs are systematically correlated with the environment (see Zaidi & Mathieson, 2020).
13. I am grateful to an anonymous reviewer, whose suggestions enhanced my discussion of this particular challenge.
14. Notably, downward causation is distinct from what is known as "evocative gene–environment correlation" and "active gene–environment correlation." The former is the term for genetic propensities evoking environmental responses (e.g., a pugilistic person evokes hostility from others), whereas the latter refers to individuals' genetically influenced propensities selecting them into specific environments (e.g., a pugilistic person takes boxing classes). Downward causation, by contrast, refers to social forces acting on (selecting and sorting) individuals based on phenotypes. See Appendix A.3 for an elaborated discussion.
15. Notably, even expansively defined risk loci may not actually contain the causal variant(s). Research using simulations or well-characterized genetic diseases demonstrates that low-frequency causal variants can generate GWAS signals that extend over *millions* of base pairs and numerous haplotypes in what is known as "long range LD" (Dickson, Wang, Krantz, Hakonarson, & Goldstein, 2010).
16. Genes in risk loci may be several or zero, and there is often no direct link to specific genes despite the use of "genes for" language that implies otherwise (e.g., "mothers with more *education-related genes* are generally healthier and more financially stable during pregnancy"; Armstrong-Carter et al., 2020; emphasis added).
17. This context-dependency reflects the social reality of these "traits" and behaviors, which I have argued, following others, makes them unsuited for to a genetic reductionist epistemology (see e.g., Burt, 2023; also Dupré, 2012; Lewontin, Rose, & Kamin, 1984; Richardson, 2017).
18. Although ethical considerations are not our focus, I question the notion of targeting interventions to those who might need extra support because of high genetic risk vs. those whose performances or whose teacher evaluations indicate they are at high risk, for whatever reason. Moreover, the use of PGSs as indicators of potential raises a host of ethical concerns, including stigma and self-limiting perceptions of one's potential.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

19. To this, some may respond that social scientists should be able to explore whatever outcomes they like and, even if not socially important, the findings "advance science." Perhaps, but I don't see scholars studying the genetic architecture of whether people have "ever eaten sushi," "ever played golf," or "only engage in sex in the missionary position in one's bed."
20. Although what is non-trivial is not always clear. Studies employing PGSs that explain ~1% or less of the variance in some outcome have been framed as non-trivial (Mills et al., 2018).
21. I would also note that from the fact that I am holding sociogenomics to a rigorous scientific standard, it does not follow that I do not believe that standard social science models should not be rigorous. That said, there is, in my view, a qualitative difference in promoting the view of partial, environmentally confounded PGSs as fixed genetic indicators of innate potential and using partial measures of socioeconomic status on complex social outcomes for several reasons that are, unfortunately, out of scope.
22. In addition to nuclear DNA, we have mitochondrial DNA (mtDNA) – a relatively tiny, maternally inherited, circular DNA molecule containing 37 genes. Unless otherwise noted, my discussions refer to nuclear DNA.
23. The number of human genes is continually updated (revised up and down) and varies across official counts because of slight differences in definitions of genes but has stabilized around 20,000. The number can never be an exact one given variation.
24. Geneticists are moving away from the SNP to SNV distinction given the somewhat arbitrary classification and different usages of the term across disciplines. Instead, there is a move toward classifying SNVs as common (>5%), low frequency (0.5–5%), and rare (<0.5%) (Strachan & Read, 2018). However, given that the GWAS field uses the term SNP, I will do so here.
25. This is basic illustration showing processes of downward causation. As noted, most GWASs imperfectly control for ancestral differences (continental ancestry) and population substructure. However, as noted in the text downward causation is pervasive – e.g., social selection on attractiveness, height, weight, colorism – with most such factors imperfectly controlled, if controlled at all.

References

Abdellaoui, A., Verweij, K. J., & Nivard, M. G. (2022). Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature Genetics*, *54*(9), 1345–1354.

Armstrong-Carter, E., Trejo, S., Hill, L. J., Crossley, K. L., Mason, D., & Domingue, B. W. (2020). The earliest origins of genetic nurture: The prenatal environment mediates the association between maternal genetics and child development. *Psychological Science*, *31*(7), 781–791.

Backman, J. D., Li, A. H., Marcketta, A., Sun, D., Mbatchou, J., Kessler, M. D., … Balasubramanian, S. (2021). Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*, *599*(7886), 628–634.

Barban, N., Jansen, R., De Vlaming, R., Vaez, A., Mandemakers, J. J., Tropf, F. C., … Nolte, I. M. (2016). Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature Genetics*, *48*(12), 1462–1472.

Barth, D., Papageorge, N. W., & Thom, K. (2020). Genetic endowments and wealth inequality. *Journal of Political Economy*, *128*(4), 1474–1522.

Barton, N., Hermisson, J., & Nordborg, M. (2019). Population genetics: Why structure matters. *eLife*, *8*, e45380.

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., … Herd, P. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, *115*(31), E7275–E7284.

Belsky, D. W., & Harden, K. P. (2019). Phenotypic annotation: Using polygenic scores to translate discoveries from genome-wide association studies from the top down. *Current Directions in Psychological Science*, *28*(1), 82–90. doi: 10.1177/0963721418807729

Belsky, D. W., & Israel, S. (2014). Integrating genetics and social science: Genetic risk scores. *Biodemography and Social Biology*, *60*(2), 137–155.

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., … Williams, B. S. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychological Science*, *27*(7), 957–972.

Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., … Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *eLife*, *8*, e39725. doi: 10.7554/elife.39725

Bliss, C. (2018). *Social by nature: The promise and peril of sociogenomics*. Stanford University Press.

Boardman, J. D., & Fletcher, J. M. (2021). Evaluating the continued integration of genetics into medical sociology. *Journal of Health and Social Behavior*, *62*(3), 404–418.

Bochukova, E. G., Huang, N., Keogh, J., Henning, E., Purmann, C., Blaszczyk, K., … O'Rahilly, S. (2010). Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature*, *463*(7281), 666–670.

Bolyard, A., & Savelyev, P. A. (2020). Understanding the education polygenic score and its interactions with SES in determining health in young adulthood. *Available at SSRN* 3397735.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*(7), 1177–1186. doi: 10.1016/j.cell.2017.05.038

Braudt, D. B. (2018). Sociogenomics in the 21st century: An introduction to the history and potential of genetically informed social science. *Sociology Compass*, *12*(10), e12626.

Browning, S. R., & Browning, B. L. (2011). Population structure can inflate SNP-based heritability estimates. *The American Journal of Human Genetics*, *89*(1), 191–193.

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. Å., … Davies, N. M. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nature Communications*, *11*(1), 3519. doi: 10.1038/s41467-020-17117-4

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., … Neale, B. M. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*(3), 291–295. doi: 10.1038/ng.3211

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., … Sollis, E. (2019). The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47* (D1), D1005–D1012.

Burt, C. H. (2018). Racial discrimination and cultural adaptations: An evolutionary developmental approach. *Advances in Criminological Theory: Building a Black Criminology*, *24*, 207–252.

Burt, C. H. (2022). Cultural evolutionary theory is not enough: Vague culture, neglect of structure, and the absence of theory in behavior genetics. *Behavioral and Brain Sciences*, *45*, E157. doi: 10.1017/S0140525X21001552

Burt, C. H. (2023). Irreducibly social: Why biocriminology's ontoepistemology is incompatible with the social reality of crime. *Theoretical Criminology*, *27*(1), 85–104.

Burt, C. H., & Munafò, M. (2021). Has GWAS lost its status as a paragon of open science? *PLoS Biology*, *19*(5), e3001242.

Bycroft, C., Fernandez-Rozadilla, C., Ruiz-Ponte, C., Quintela, I., Carracedo, Á, Donnelly, P., & Myers, S. (2019). Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula. *Nature Communications*, *10*(1), 1–14.

Byrne, R. P., van Rheenen, W., van den Berg, L. H., Veldink, J. H., & McLaughlin, R. L. (2020). Dutch population structure across space, time and GWAS design. *Nature Communications*, *11*(1), 1–11.

Cardon, L. R., & Palmer, L. J. (2003). Population stratification and spurious allelic association. *The Lancet*, *361*(9357), 598–604. doi: 10.1016/s0140-6736(03)12520-2

Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics*, *16*(3), 319–341.

Cesarini, D., & Visscher, P. M. (2017). Genetics and educational attainment. *NPJ Science of Learning*, *2*(1), 1–7.

Charney, E. (2022). The "Golden Age" of behavior genetics? *Perspectives on Psychological Science*, *17*(4), 1188–1210.

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., Mcadams, T. A., … Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science*, *31*(5), 582–591. doi: 10.1177/0956797620904450

Chiang, C., Scott, A. J., Davis, J. R., Tsang, E. K., Li, X., Kim, Y., … Montgomery, S. B. (2017). The impact of structural variation on human gene expression. *Nature Genetics*, *49*(5), 692–699.

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

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., … Wang, H. (2020). A structural variation reference for medical and population genetics. *Nature*, *581*(7809), 444–451.

Conley, D. (2016). Socio-genomic research using genome-wide molecular data. *Annual Review of Sociology*, *42*, 275–299.

Conley, D., & Fletcher, J. (2017). *The genome factor*. Princeton University Press.

Coop, G., & Przeworski, M. (2022). Lottery, luck, or legacy. A review of "The Genetic Lottery: Why DNA matters for social equality" *Evolution*, *76*(4), 846–853.

Crouch, D. J., & Bodmer, W. F. (2020). Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. *Proceedings of the National Academy of Sciences*, *117*(32), 18924–18933.

Curtis, D. (2018). Polygenic risk score for schizophrenia is more strongly associated with ancestry than with schizophrenia. *Psychiatric Genetics*, *28*(5), 85–89.

Dandine-Roulland, C., Bellenguez, C., Debette, S., Amouyel, P., Génin, E., & Perdry, H. (2016). Accuracy of heritability estimations in presence of hidden population stratification. *Scientific Reports*, *6*, 26471.

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics*, *55* (4), 997–1004. doi: 10.1111/j.0006-341x.1999.00997.x

Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., … Sher, K. J. (2015). Candidate gene–environment interaction research: Reflections and recommendations. *Perspectives on Psychological Science*, *10*(1), 37–59.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Dickson, S. P., Wang, K., Krantz, I., Hakonarson, H., & Goldstein, D. B. (2010). Rare variants create synthetic genome-wide associations. *PLoS Biology*, *8*(1), e1000294.
Domingue, B., Trejo, S., Armstrong-Carter, E., & Tucker-Drob, E. (2020). Interactions between polygenic scores and environments: Methodological and conceptual challenges.
Doolittle, W. F. (2013). Is junk DNA bunk? A critique of ENCODE. *Proceedings of the National Academy of Sciences*, *110*(14), 5294–5300.
Dupré, J. (2012). *Processes of life: Essays in the philosophy of biology*. Oxford University Press.
Duster, T. (2015). A post-genomic surprise. The molecular reinscription of race in science, law and medicine. *The British Journal of Sociology*, *66*(1), 1–27.
Edwards, S. L., Beesley, J., French, J. D., & Dunning, A. M. (2013). Beyond GWASs: Illuminating the dark road from association to function. *The American Journal of Human Genetics*, *93*(5), 779–797.
Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicius, V., … Volk, A. A. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, *48*(3), 598.
ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia of DNA elements) project. *Science (New York, N.Y.)*, *306*(5696), 636–640.
Evans, D. M., Visscher, P. M., & Wray, N. R. (2009). Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Human Molecular Genetics*, *18*(18), 3525–3531.
Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. *Science*, *190*(4220), 1163–1168.
Feldman, M. W., Lewontin, R. C., & King, M.-C. (2003). Race: A genetic melting-pot. *Nature*, *424*(6947), 374–374.
Freese, J. (2008). Genetics and the social science explanation of individual outcomes. *American Journal of Sociology*, *114*(S1), S1–S35.
Freese, J. (2018). The arrival of social science genomics. *Contemporary Sociology*, *47*(5), 524–536.
Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American Journal of Epidemiology*, *186*(9), 1026–1034.
Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., … Mittnik, A. (2016). The genetic history of ice age Europe. *Nature*, *534*(7606), 200–205.
Fu, W., O'connor, T. D., Jun, G., Kang, H. M., Abecasis, G., Leal, S. M., … Shendure, J. (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature*, *493*(7431), 216–220.
Ganna, A., Verweij, K. J., Nivard, M. G., Maier, R., Wedow, R., Busch, A. S., … Lichtenstein, P. (2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science*, *365*(6456), eaat7693.
Germain, P.-L., Ratti, E., & Boem, F. (2014). Junk or functional DNA? ENCODE and the function controversy. *Biology & Philosophy*, *29*(6), 807–831. doi: 10.1007/s10539-014-9441-3
Ginger, R. S., Askew, S. E., Ogborne, R. M., Wilson, S., Ferdinando, D., Dadd, T., … Green, M. R. (2008). SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium–calcium exchange activity that regulates human epidermal melanogenesis. *Journal of Biological Chemistry*, *283*(9), 5486–5495. doi: 10.1074/jbc.m707521200
Hamer, D. H. (2000). Beware the chopsticks gene. *Molecular Psychiatry*, *5*(1), 11–13.
Harden, K. P. (2021a). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.
Harden, K. P. (2021b). "Reports of my death were greatly exaggerated": Behavior genetics in the postgenomic era. *Annual Review of Psychology*, *72*, 37–60.
Harden, K. P., Domingue, B. W., Belsky, D. W., Boardman, J. D., Crosnoe, R., Malanchini, M., … Harris, K. M. (2020). Genetic associations with mathematics tracking and persistence in secondary school. *NPJ Science of Learning*, *5*(1), 1–8.
Harden, K. P., & Koellinger, P. D. (2020). Using genetics for social science. *Nature Human Behaviour*, *4*(6), 567–576.
Hart, S. A., Little, C., & van Bergen, E. (2021). Nurture might be nature: Cautionary tales and proposed solutions. *NPJ Science of Learning*, *6*(1), 1–12.
Haworth, S., Mitchell, R., Corbin, L., Wade, K. H., Dudding, T., Budu-Aggrey, A., … Smith, G. D. (2019). Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications*, *10*(1), 1–9.
Healey, M. D., & Ellis, B. J. (2007). Birth order, conscientiousness, and openness to experience: Tests of the family-niche model of personality using a within-family methodology. *Evolution and Human Behavior*, *28*(1), 55–59.
Herd, P., Freese, J., Sicinski, K., Domingue, B. W., Mullan Harris, K., Wei, C., & Hauser, R. M. (2019). Genes, gender inequality, and educational attainment. *American Sociological Review*, *84*(6), 1069–1098.
Herd, P., Mills, M. C., & Dowd, J. B. (2021). Reconstructing sociogenomics research: Dismantling biological race and genetic essentialism narratives. *Journal of Health and Social Behavior*, *62*(3), 419–435.
Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., … Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*, 1. doi: 10.1038/s41467-019-13585-5
Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., … van der Zee, M. D. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592.
Hubbard, R. (1999). *Exploding the gene myth: How genetic information is produced and manipulated by scientists, physicians, employers, insurance companies, educators, and law enforcers*. Beacon Press.
Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., . . . Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America* (pp. 517–523). Basic Books.
Jensen, A. R. (1967). How much can we boost IQ and scholastic achievement?.
Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., … Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, *42*(4), 348–354. doi: 10.1038/ng.548
Kaplan, J. M., & Turkheimer, E. (2021). Galton's quincunx: Probabilistic causation in developmental behavior genetics. *Studies in History and Philosophy of Science Part A*, *88*, 60–69.
Karakachoff, M., Duforet-Frebourg, N., Simonet, F., Le Scouarnec, S., Pellen, N., Lecointe, S., … Froguel, P. (2015). Fine-scale human genetic structure in western France. *European Journal of Human Genetics*, *23*(6), 831–836.
Kerminen, S., Havulinna, A. S., Hellenthal, G., Martin, A. R., Sarin, A.-P., Perola, M., … Ripatti, S. (2017). Fine-scale genetic structure in Finland. *G3: Genes, Genomes, Genetics*, *7*(10), 3459–3468.
Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Masson, G. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428.
Kweon, H., Burik, C., Karlsson Linnér, R., De Vlaming, R., Okbay, A., Martschenko, D., … Koellinger, P. (2020). Genetic fortune: Winning or losing education, income, and health.
Laird, N. M., & Lange, C. (2006). Family-based designs in the age of large-scale gene-association studies. *Nature Reviews Genetics*, *7*(5), 385–394.
Lamason, R. L., Mohideen, M.-A. P., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., … Humbert, J. E. (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, *310*(5755), 1782–1786.
Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, *265* (5181), 2037–2048.
Lawson, D. J., Davies, N. M., Haworth, S., Ashraf, B., Howe, L., Crawford, A., … Timpson, N. J. (2020). Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? *Human Genetics*, *139*(1), 23–41. doi: 10.1007/s00439-019-02014-8
Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Linnér, R. K. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112–1121.
Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., … Lawson, D. J. (2015). The fine-scale genetic structure of the British population. *Nature*, *519*(7543), 309–314.
Lewontin, R. C. (1974). Annotation: The analysis of variance and the analysis of causes. *American Journal of Human Genetics*, *26*(3), 400.
Lewontin, R. C., Rose, S., & Kamin, L. J. (1984). *Not in our genes*. Pantheon Books.
Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., … Kenny, E. E. (2017). Human demographic history impacts genetic risk prediction across diverse populations. *The American Journal of Human Genetics*, *100*(4), 635–649.
Martschenko, D., Trejo, S., & Domingue, B. W. (2019). Genetics and education: Recent developments in the context of an ugly history and an uncertain future. *AERA Open*, *5*(1), 233285841881051. doi: 10.1177/2332858418810516
Mathieson, I., & Mcvean, G. (2012). Differential confounding of rare and common variants in spatially structured populations. *Nature Genetics*, *44*(3), 243–246. doi: 10.1038/ng.1074
McClellan, J., & King, M.-C. (2010). Genetic heterogeneity in human disease. *Cell*, *141*(2), 210–217.
Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829.
Mills, M. C., Barban, N., & Tropf, F. C. (2018). The sociogenomics of polygenic scores of reproductive behavior and their relationship to other fertility traits. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *4*(4), 122–136.
Mills, M. C., & Tropf, F. C. (2020). Sociology, genetics, and the coming of age of sociogenomics. *Annual Review of Sociology*, *46*, 553–581.
Monk Jr, E. P., Esposito, M. H., & Lee, H. (2021). Beholding inequality: Race, gender, and returns to physical attractiveness in the United States. *American Journal of Sociology*, *127*(1), 194–241.
Morris, T. T., Davies, N. M., Hemani, G., & Smith, G. D. (2020a). Population phenomena inflate genetic associations of complex social traits. *Science Advances*, *6*(16), eaay0328.
Morris, T. T., Davies, N. M., & Smith, G. D. (2020b). Can education be personalised using pupils' genetic data? *eLife*, *9*, e49962.
Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J. K., & Przeworski, M. (2020). Variable prediction accuracy of polygenic scores within an ancestry group. *eLife*, *9*, e48376. doi: 10.7554/elife.48376
Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science*, *310*(5746), 321–324.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Novembre, J., & Barton, N. H. (2018). Tread lightly interpreting polygenic tests of selection. *Genetics*, *208*(4), 1351–1355.
O'Connor, T. D., Fu, W., Mychaleckyj, J. C., Logsdon, B., Auer, P., Carlson, C. S., … Akey, J. M. (2015). Rare variation facilitates inferences of fine-scale population structure in humans. *Molecular Biology and Evolution*, *32*(3), 653–660. doi: 10.1093/molbev/msu326
Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., … Gjorgjieva, T. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449.
Panofsky, A. (2014). *Misbehaving science: Controversy and the development of behavior genetics*. University of Chicago Press.
Pennisi, E. (2012). ENCODE project writes eulogy for junk DNA. *Science*, *337*(6099), 1159–1161.
Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., … Brick, L. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell*, *179*(3), 589–603.
Plomin, R. (2019). *Blueprint: How DNA makes us who we are*. MIT Press.
Plomin, R., & Von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*(3), 148–159. doi: 10.1038/nrg.2017.104
Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. doi: 10.1038/ng1847
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Daly, M. J. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.
Richardson, K. (2017). *Genes, brains, and human potential*. Columbia University Press.
Richardson, K., & Jones, M. C. (2019). Why genome-wide associations with cognitive ability measures are probably spurious. *New Ideas in Psychology*, *55*, 35–41. doi: 10.1016/j.newideapsych.2019.04.005
Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., … Agrawal, A. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *340*(6139), 1467–1471.
Ronda, V., Agerbo, E., Bleses, D., Bo Mortensen, P., Børglum, A., Hougaard, D. M., … Rosholm, M. (2020). Family disadvantage, gender, and the returns to genetic human capital. *The Scandinavian Journal of Economics*, *124*(2), 550–578.
Schaid, D. J., Chen, W., & Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, *19*(8), 491–504. doi: 10.1038/s41576-018-0016-z
Shen, H., & Feldman, M. W. (2020). Genetic nurturing, missing heritability, and causal analysis in genetic statistics. *Proceedings of the National Academy of Sciences*, *117*(41), 25646–25654. doi: 10.1073/pnas.2015869117
Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., & Waterston, R. H. (2017). DNA sequencing at 40: Past, present and future. *Nature*, *550*(7676), 345–353.
Shendure, J. A., Porreca, G. J., & Church, G. M. (2008). Overview of DNA sequencing strategies. *Current Protocols in Molecular Biology*, *81*(1), 7.1.1–7.1.11.
Simons, R. L., Burt, C. H., Barr, A. B., Lei, M. K., & Stewart, E. (2014). Incorporating routine activities, activity spaces, and situational definitions into the social schematic theory of crime. *Criminology; An interdisciplinary Journal*, *52*(4), 655–687.
Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., … Sunyaev, S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife*, *8*, e39702. doi: 10.7554/elife.39702
Strachan, T., & Read, A. P. (2018). *Human molecular genetics*. Garland.
Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., … Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, *526*(7571), 75–81. doi: 10.1038/nature15394
Sulloway, F. J. (2001). Birth order, sibling competition, and human behavior. In *Conceptual challenges in evolutionary psychology* (pp. 39–83). Springer.
Takumi, T., & Tamada, K. (2018). CNV biology in neurodevelopmental disorders. *Current Opinion in Neurobiology*, *48*, 183–192.
Taliun, D., Harris, D. N., Kessler, M. D., Carlson, J., Szpiech, Z. A., Torres, R., … Kang, H. M. (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *Nature*, *590*(7845), 290–299.
Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, *20*(8), 467–484. doi: 10.1038/s41576-019-0127-1
Telenti, A., Pierce, L. C. T., Biggs, W. H., Di Iulio, J., Wong, E. H. M., Fabani, M. M., … Venter, J. C. (2016). Deep sequencing of 10,000 human genomes. *Proceedings of the National Academy of Sciences*, *113*(42), 11901–11906. doi: 10.1073/pnas.1613365113
The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. doi: 10.1038/nature15393
The International HapMap Consortium (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320. doi: 10.1038/nature04226
Trejo, S., Belsky, D. W., Boardman, J. D., Freese, J., Harris, K. M., Herd, P., … Domingue, B. W. (2018). Schools as moderators of genetic associations with life course attainments: Evidence from the WLS and Add heath. *Sociological Science*, *5*, 513–540.
Trejo, S., & Domingue, B. W. (2019). Genetic nature or genetic nurture? Quantifying bias in analyses using polygenic scores. *bioRxiv*, 524850.
Turkheimer, E. (2016). Weak genetic explanation 20 years later: Reply to Plomin et al. (2016). *Perspectives on Psychological Science*, *11*(1), 24–28.
Wedow, R., Zacher, M., Huibregtse, B. M., Mullan Harris, K., Domingue, B. W., & Boardman, J. D. (2018). Education, smoking, and cohort change: Forwarding a multidimensional theory of the environmental moderation of genetic effects. *American Sociological Review*, *83*(4), 802–832.
Wertz, J., Belsky, J., Moffitt, T. E., Belsky, D. W., Harrington, H., Avinun, R., … Caspi, A. (2019). Genetics of nurture: A test of the hypothesis that parents' genetics predict their observed caregiving. *Developmental Psychology*, *55*(7), 1461.
Wertz, J., Caspi, A., Belsky, D. W., Beckley, A. L., Arseneault, L., Barnes, J., … Morgan, N. (2018). Genetics and crime: Integrating new genomic discoveries into psychological research about antisocial behavior. *Psychological Science*, *29*(5), 791–803.
Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, *26*(17), 2190–2191.
Willer, C. J., Speliotes, E. K., Loos, R. J., Li, S., Lindgren, C. M., Heid, I. M., … Lamina, C. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature Genetics*, *41*(1), 25.
Wray, N. R., Goddard, M. E., & Visscher, P. M. (2007). Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Research*, *17*(10), 1520–1528.
Wright, J. P., & Cullen, F. T. (2012). The future of biosocial criminology: Beyond scholars' professional ideology. *Journal of Contemporary Criminal Justice*, *28*(3), 237–253.
Wu, Y., Zhong, X., Lin, Y., Zhao, Z., Chen, J., Zheng, B., … Lu, Q. (2021). Estimating genetic nurture with summary statistics of multigenerational genome-wide association studies. *Proceedings of the National Academy of Sciences*, *118*(25), e2023184118.
Yong, E. (2019). A waste of 1,000 research papers. *The Atlantic*. Retrieved from https://www.theatlantic.com/science/archive/2019/05/waste-1000-studies/589684/
Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., … Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*(9), 1304–1310.
Zaidi, A. A., & Mathieson, I. (2020). Demographic history mediates the effect of stratification on polygenic scores. *eLife*, *9*, e61548. doi: 10.7554/elife.61548
Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, *109*(4), 1193–1198.

Appendix A

In what follows, I provide a concise overview of the genomics of sociogenomics, including an introduction to genomics, the types of genetic variation, and their potential effects. This discussion is necessarily abbreviated and detailed as "all going well" (e.g., chromosomal aneuploidies are not discussed). This is followed by a short elaboration of downward causation and artificial genetic signals and a comparison with "authentic" genetic signals and conditional genetic effects.

A.1 Basic genetics of sociogenomics

(Nuclear) DNA are the focus of human genetics.[22] Humans have 46 chromosomes, each of which is a very long double-stranded molecule of DNA arranged in the famous double helix. We inherit 22 matching pairs of non-sex chromosomes, one each from our mothers and fathers. In addition, each of us inherits an X chromosome from our mother and either an X or Y chromosome from our father that determines sex, all going well. Each chromosome is composed of a linear sequence of *nucleotides* – the building blocks of DNA. Nucleotides are composed of three parts: a deoxyribose sugar, a phosphate group, and one of four nucleic acid bases: adenine (A), thymine (T), guanine (G), and cytosine (C). The order of these bases on our chromosomes is our genetic code. Altogether, the human genome contains ~6 billion bases (3 billion base pairs [bp]).

Genes are sequences of DNA scattered on our chromosomes that serve as templates for making an RNA product (that becomes a protein or functional RNA product with subsequent processing). The canonical gene is a protein-coding gene – a stretch of DNA that encodes the sequence of amino acids that will be folded into a functional protein. So-called "non-coding RNA genes" are DNA sequences that encode functional RNA products, which perform essential cellular functions, including facilitating and regulating gene

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

expression. Following others, when I use "gene," I refer to protein-coding genes.

Our DNA are informational storage molecules. Like recipes, genes are not self-activating but are used by cellular machinery to create proteins via coordinated cellular mechanisms, especially RNAs and ribosomes (Hubbard, 1999). Messenger RNAs, which are specified by the "genetic [protein] code," serve as the information-transfer intermediary between DNA and proteins. The language or "ingredients" in our genetic code are three-base sequences, known as codons, which specify an amino acid (or a stop message). There are 20 amino acids and 64 codons, of which one is a "start" codon and three codons specify a stop transcription message (like a period). Each codon specifies only amino acid, but most amino acids are encoded by two or more codons, primarily because of redundancy at the third base.

Excepting male-specific genes on the Y chromosome, we inherit two copies of each gene, one from each parent. Overall, humans have ~20,000 (protein-coding) genes,[23] slightly more than chicken and fewer than half the genes of rice (~50,000 genes). Despite only having ~20,000 genes, humans can produce more than 100,000 proteins. Our complexity is not a function of our gene number (or genome size) but by complexities in gene regulation. This one gene → multiple proteins potential is facilitated by a variety of RNA-mediated mechanisms, including alternative splicing – where the same "gene" (more precisely, mRNA transcript) is "spliced" in different ways to make different amino acid chains; "readthrough" or "conjoined" genes, where two adjacent genes are transcribed together; as well as post-translational modifications, where different folding of polypeptides creates different functional proteins. In the same way a recipe does not make a cake, genes do not make a protein, much less a phenotype.

Despite getting the most attention, protein-coding DNA only comprises about 1.3% of our genome. Much of the remainder of our DNA was once thought to be largely junk; however, research revealed that most of our genome contains signals of function (ENCODE Project Consortium, 2004). How much of our genome is, in fact, functional (~5–85%) remains debated (Doolittle, 2013; Germain, Ratti, & Boem, 2014; Pennisi, 2012).

A.2 Overview of genetic variation

A.2.1 Types and consequences of genetic variation

There are three main classes of DNA variants. Almost always, GWASs examine only a subtype of the first of these.

A.2.1.1 Single-nucleotide variants (SNVs) and single-nucleotide polymorphisms (SNPs)

The first and by far the most common variant – accounting for almost 87% of all variants between people – are *single-nucleotide variants* (SNVs). An SNV exists where, for example, at specific position on the genome most people may have an A but a minority of people have a C. SNVs that are "common" occur in at least 1% (though sometimes >0.5%) of a population are known as *single-nucleotide polymorphisms* (SNPs – pronounced "snips"). SNPs are thus the subset of SNVs that are "common."[24] Most SNPs are ancient mutations that predate the out of Africa dispersal of humans some 50–100 thousand years ago and are thus shared by all human populations.

At present there are more than 475 million validated SNVs, most of which are rare. Many (roughly half) of these SNVs are "singletons"; that is, they are observed in only one individual in a sample (Taliun et al., 2021). Although most SNVs are rare (i.e., not SNPs), most (>95%) of the SNVs in an individual genome are common (are SNPs) (Taliun et al., 2021; Telenti et al., 2016). In total, there are ~10–20 million SNPs in the human genome, with variation because of how one defines "common" (The 1000 Genomes Project Consortium, 2015).

Most SNVs are bi-allelic (come in two forms), but some are tri-allelic or quad-allelic. Bi-allelic SNPs are the form of variation examined in most GWASs and used in the creation of PGSs.

A.2.1.2 (Short) insertion–deletions (indels)

A second class of variants comprises short insertions and deletions (indels), which includes duplications, deletions, or insertions up to 50 bp. (Short) copy number variants (CNVs) (including those which have a variable number of tandem unit repeats (or VNTRs), such as a sequence TTACTGC repeated 4–8 times), are included as "indels" or "delins" here as in genome-sequencing projects.

Indels are relatively common (account for ~13% of human sequence variation) and have multiple alleles leading to significant genetic heterogeneity (which is why short-sequence repeats are useful in forensic DNA testing). Indels are rarely measured in GWASs (Tam et al., 2019).

A.2.1.3 Structural variants

The remaining class of genetic variation, structural variants (SVs), is DNA rearrangements (deletions, duplications, or inversions) involving more than 50 bp. In the past SVs were defined as larger sequence changes typically up to 1 kb, but now are defined as smaller changes and include CNVs larger than 50 bp (Strachan & Read, 2018).

Although SVs are relatively uncommon (accounting for only ~0.15% of the variants, which translates to about 7,500 per genome), they account for more (nearly 2× more) overall nucleotide (sequence) differences than the two other variant types combined given their size (Collins et al., 2020; Sudmant et al., 2015). Notably, measuring SVs is much more difficult and less common given that the short-read, efficient sequencing technology that predominates does not measure SVs well (Shendure et al., 2017; Shendure, Porreca, & Church, 2008). Long-read sequencing suggests that there may be several-fold more SVs that are hidden because of systematic biases in detection (Sudmant et al., 2015).

A.3 Effects of genetic variants

Notably, most of our variants lie outside of coding regions with no known (or expected) functional impact (i.e., [putatively] "nonfunctional variants"). That said, a recent deep sequencing study observed that one-third of human protein-coding genes show some variation among individuals in the amino acid sequences they encode (Taliun et al., 2021). As discussed in the text, functional variants either alter gene product (e.g., the protein produced) or gene dosage (e.g., the amount of protein produced).

SNVs are classified by their functional effects in coding regions. "Synonymous" SNVs are non-functional base changes that do not alter the amino acid and protein product, whereas "non-synonymous" SNVs are those that change the amino acid sequence. There are three types of non-synonymous SNVs: missense, nonsense, and read-through variants. Missense variants change the amino acid (e.g., CCU → ACU would change the amino acid from proline to threonine) and can have significant to no noticeable effect on the protein and its efficacy (think switching sugar with pepper in a recipe vs. switching onion powder with garlic powder). Nonsense mutations cause a premature stop codon (e.g., GGA [glycine] → UGA [stop]). These effects tend to be more significant than missense changes, much like a recipe that just ended randomly early. Finally, read-through or nonstop mutations change a stop codon to an amino acid codon, causing the polypeptide to be longer than it should be (e.g., UGA [stop] → GGA [glycine]), akin to just adding more ingredients to a recipe.

Unlike SNVs, indels and SVs affect more than 1 base pair and thus produce differences in the lengths of DNA sequences across people. These variants can have significant functional consequences given they alter more sequences and can result in coding frameshifts, which refer to shifts in the entire coding sequence which can markedly alter the composition of the resulting polypeptide product. A useful analogy to frameshift effects is the removal of a few letters from a sentence. For example, deleting a few letters in the first sentence in the statement: "I am going to the store tomorrow. Is that okay?" makes the sentence gobbledygook: "I am gothe st oreto morrowistha."

A.4 Meaning of downward and upward causation in a genetic context

As, I discuss in the main text, the counterfactual "variant substitution effect" model underlying GWASs and PGSs cannot distinguish between authentic

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

genetic associations and artificial ones representing downward causation. In GWASs and thus PGSs, both signals are identified as causal.

Authentic genetic variants are those that act in biological pathways shaping traits or diseases, such as variants affecting age-related macular degeneration or Huntington's disease. In these cases, variants causally influence phenotypes through biological pathways (e.g., via non-synonymous substitutions causing amino acid replacement). By contrast, downward causation refers to the situation where socio-environmental forces are the causal forces driving a genotype–phenotype association. Downward causation is "downward" because social forces are acting (down) on traits or other differences, which are shaped by genetic differences (thereby generating observed genetic associations). In these cases, identified genetic differences are not causally involved in the biology of trait or behavior differences; the signals are artificial because they reflect social not genetic processes.

For a real-world example of downward causation, African Americans were excluded from many educational institutions before and during Jim Crow on the basis of their race (and of course differentially admitted even after Jim Crow due to persisting discrimination). In this case, (racist) social structures acted upon 'racialized' genetic differences, such as alleles related to skin pigmentation, to exclude or restrict individuals for reasons biologically unrelated to educational attainment. In a GWAS,[25] such alleles would be identified as causing differences in educational attainment, but these association signals would, of course, be artificial.

Notably, downward causation is distinct from (causal) conditional genetic effects, in which genetic differences influence phenotypes (through biological pathways) only in some context. Conditional genetic effects are causally biologically involved in trait differences, whereas genetic variants reflecting downward causation are not.

Finally, the distinction between downward causation and an authentic genetic influence is not normative one. The distinction reflects the direction of causality and the relevance of the genetic difference to the biology of the trait, whether or not we think such differences are fair or just.

Open Peer Commentary

Don't miss the chance to reap the fruits of recent advances in behavioral genetics

Nina Alexander[a,b], Sabrina Illius[c,d], Dennis Feyerabend[e], Jan Wacker[f] and Ulf Liszkowski[e]

[a]Department of Psychiatry and Psychotherapy, Philipps University Marburg, Marburg, Germany; [b]Center for Mind, Brain and Behavior, Philipps University Marburg, Marburg, Germany; [c]Department of Psychology, Faculty of Human Sciences, Medical School Hamburg, Hamburg, Germany; [d]ICAN Institute for Cognitive and Affective Neuroscience, Medical School Hamburg, Hamburg, Germany; [e]Department of Developmental Psychology, University of Hamburg, Hamburg, Germany and [f]Department of Differential Psychology and Psychological Assessment, University of Hamburg, Hamburg, Germany
nina.alexander@uni-marburg.de; UKGM Gießen/Marburg-Team
sabrina.illius@medicalschool-hamburg.de
Dennis.Feyerabend@uni-hamburg.de
jan.wacker@uni-hamburg.de
ulf.liszkowski@uni-hamburg.de; https://www.koku.uni-hamburg.de/en/koku-team/liszkowski.html

doi:10.1017/S0140525X22002497, e208

Abstract

In her target article, Burt revives a by now ancient debate on nature and nurture, and the ways to measure, disentangle, and ultimately trust one or the other of these forces. Unfortunately, she largely dismisses recent advances in behavior genetics and its huge potential in contributing to a better *prediction* and *understanding* of complex traits in social sciences.

In the light of mean heritability estimates of 49% across 17,804 traits that derive from 14,558,903 twin pairs of different cultures (Polderman et al., 2015), we argue that the question is not *whether* but rather *how* to integrate genetic data to advance our understanding of human psychology. Owing to the unprecedented advances of massive parallel sequencing, large-scale genome-wide association studies (GWASs) have become increasingly accessible and affordable in social sciences. Accordingly, the predictive power of polygenic scores (PGSs) is steadily rising proportional to the GWAS sample size (Mitchell et al., 2021) and can already explain a substantial amount of variance in complex phenotypes such as educational attainment (~12–16%, Okbay et al., 2022) or externalizing traits (~10%, Karlsson Linnér et al., 2021). Moreover, studies applying a multi-PGS approach suggest that predictive accuracy for a given outcome can be further improved by combining PGSs of different traits (Allegrini et al., 2019; Krapohl et al., 2018). Effect sizes of some single, well-performing PGSs are already comparable to those achieved by conventional measures used in social sciences such as family characteristics (Derzon, 2010) and neighborhood disadvantage (Winslow & Shaw, 2007). This upward trend is expected to continue because of the steady progress in discovering rare genetic variants underlying complex trait heritability that are still insufficiently tagged by current GWASs (Dou et al., 2021). Estimates from large whole-genome sequencing data sets identified rare variants as a major source of the discrepancy between single-nucleotide polymorphism (SNP)-based and pedigree estimates of heritability for complex, polygenic traits such as height (Wainschtein et al., 2022). In contrast, the frequently discussed concern that PGSs of complex traits are doomed to miss a substantial amount of non-additive variance is currently not well supported. Instead, average estimates from large samples of unrelated individuals suggest that dominance effects explain at most a very small amount of variance in complex traits (Hivert et al., 2021; Okbay et al., 2022). Consequently, it is only legitimate to assume that PGSs are just about to unfold their full predictive potential.

Burt is further concerned that PGSs are inevitably compromised by environmental confounding, whereas others argue that traditional environmental measures, for example childhood maltreatment, are also confounded by substantial heritable components (Dalvie et al., 2020; Hart, Little, & van Bergen, 2021; Warrier et al., 2021). As Burt rightly cautions, quantifying the extent by which the predictive power of PGSs results from genotype–environment correlation (rGE) is challenging, but indeed essential for their adequate *interpretation*. Large within-family studies (e.g., using parent–offspring trios) have significantly contributed to more precise estimates of rGE (Chen et al., 2022) and could be further advanced through developmental approaches starting from infancy when environmental variance is still reduced (Falck-Ytter et al., 2021). Importantly, however, disentangling direct from indirect genetic effect of PGS is less relevant

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

whenever the primary goal is to improve (risk) *prediction* accuracy, given that rGE does not undermine a PGS's predictive capacities (see Plomin & von Stumm, 2022). Moreover, even those PGSs where a relatively large amount of predictive power is not derived from direct genetic effects (e.g., educational attainment, Okbay et al., 2022) still capture variance that is substantially independent of and thus incremental to the effects of traditional environmental measures, like socioeconomic status (Judd et al., 2020). Consequently, algorithms that jointly model the effects of PGSs and environmental measures are performing significantly better in predicting, for example, health outcomes (Adeyemo et al., 2021; Martikainen et al., 2021; Østergaard et al., 2020) and cognitive functioning (von Stumm et al., 2020) compared to those that include traditional non-genetic measures only. The potential to improve prediction by combining genetic and environmental data even translates down to epigenetic modifications, which are increasingly recognized in social science studies as a potential mechanism of how life events get under the skin. Epigenome-wide analyses across independent cohorts revealed that variation in DNA methylation is best explained by additive effects and the interaction of genes and environmental forces, but almost never by environmental adversity alone (Czamara et al., 2021).

Beyond our defense of the immediate practical utility of PGSs for maximizing trait prediction, we also do not share Burt's skepticism regarding the limited potential of PGSs for advancing our etiological understanding of complex traits. The growing number of studies combining PGSs with neuroimaging, proteomic, or other multi-omic data have already provided unique insights into specific mechanisms through which polygenic predispositions exert their effects on complex phenotypes. Exemplary findings from neuroimaging studies include the identification of structural brain changes associated with PGSs for neuroticism (Opel et al., 2020) and educational attainment (Elliott et al., 2019), that, in the latter example, partly mediated the association between participants' PGS and their cognitive test performance. Moreover, PGSs have already been successfully applied to study the causal biology of complex traits, for example, in terms of identifying specific proteins underlying cardiometabolic diseases (Ritchie et al., 2021). Another promising new method to advance etiological understanding of complex traits is to construct PGSs based on gene transcription profiles targeting specific biological systems, including PGSs capturing neurotransmitter signaling pathways (Miguel et al., 2019; Restrepo-Lozano et al., 2022), immuno-metabolic markers (Kappelmann et al., 2021), or cellular stress responses (Arloth et al., 2015). For example, a recent study reported that a PGS based on corticolimbic-specific DCC gene co-expression, which modulates maturation of dopamine networks, is a better predictor of impulsivity-related phenotypes than conventional PGSs (Restrepo-Lozano et al., 2022).

To conclude, we argue that despite their indisputable limitations, PGSs hold great potential for both better *prediction* and *understanding* of complex traits in social science. Raw SNP data from genome-wide arrays can now be generated for only ~US $35 per individual test with an excellent accuracy that outcompetes those of most environmental measures (genotype concordance >98%, Hong et al., 2012). Once obtained, SNP data allow for an automated generation and flexible adaption of multiple PGSs at any time in life because of their inherent intraindividual stability. The initial struggle of identifying causal genetic variants for complex traits should not discourage us from embracing the remarkable achievements recently made in molecular behavior genetics.

Financial support. The research was funded by the State Research Funding "Hamburger Landesforschungsförderung; LFF-FV79, Project 3, Differential Ontogeny of Human Interaction" to principal investigators Liszkowski, Alexander, and Wacker.

Competing Interest. None.

References

Adeyemo, A., Balaconis, M. K., Darnes, D. R., Fatumo, S., Granados Moreno, P., Hodonsky, C. J., … Polygenic Risk Score Task Force of the International Common Disease, A. (2021). Responsible use of polygenic risk scores in the clinic: Potential benefits, risks and gaps. *Nature Medicine*, *27*(11), 1876–1884. doi: 10.1038/s41591-021-01549-6

Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry*, *24*(6), 819–827. doi: 10.1038/s41380-019-0394-4

Arloth, J., Bogdan, R., Weber, P., Frishman, G., Menke, A., Wagner, K. V., … Binder, E. B. (2015). Genetic differences in the immediate transcriptome response to stress predict risk-related brain function and psychiatric disorders. *Neuron*, *86*(5), 1189–1202. doi: 10.1016/j.neuron.2015.05.034

Chen, C., Lu, Y., Lundström, S., Larsson, H., Lichtenstein, P., & Pettersson, E. (2022). Associations between psychiatric polygenic risk scores and general and specific psychopathology symptoms in childhood and adolescence between and within dizygotic twin pairs. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *63*(12), 1513–1522. doi: 10.1111/jcpp.13605.

Czamara, D., Tissink, E., Tuhkanen, J., Martins, J., Awaloff, Y., Drake, A. J., … Binder, E. B. (2021). Combined effects of genotype and childhood adversity shape variability of DNA methylation across age. *Translational Psychiatry*, *11*(1), 88. doi: 10.1038/s41398-020-01147-z

Dalvie, S., Maihofer, A. X., Coleman, J. R. I., Bradley, B., Breen, G., Brick, L. A., … Nievergelt, C. M. (2020). Genomic influences on self-reported childhood maltreatment. *Translational Psychiatry*, *10*(1), 38. doi: 10.1038/s41398-020-0706-0

Derzon, J. H. (2010). The correspondence of family features with problem, aggressive, criminal, and violent behavior: A meta-analysis. *Journal of Experimental Criminology*, *6*(3), 263–292. doi: 10.1007/s11292-010-9098-0

Dou, J., Wu, D., Ding, L., Wang, K., Jiang, M., Chai, X., … Wang, C. (2021). Using off-target data from whole-exome sequencing to improve genotyping accuracy, association analysis and polygenic risk prediction. *Briefings in Bioinformatics*, *22*(3), bbaa084. doi: 10.1093/bib/bbaa084

Elliott, M. L., Belsky, D. W., Anderson, K., Corcoran, D. L., Ge, T., Knodt, A., … Hariri, A. R. (2019). A polygenic score for higher educational attainment is associated with larger brains. *Cerebral Cortex (New York, N.Y.: 1991)*, *29*(8), 3496–3504. doi: 10.1093/cercor/bhy219

Falck-Ytter, T., Hamrefors, L., Siqueiros Sanches, M., Portugal, A. M., Taylor, M., Li, D., … Ronald, A. (2021). The Babytwins Study Sweden (BATSS): A multi-method infant twin study of genetic and environmental factors influencing infant brain and behavioral development. *Twin Research and Human Genetics*, *24*(4), 217–227. doi: 10.1017/thg.2021.34

Hart, S. A., Little, C., & van Bergen, E. (2021). Nurture might be nature: Cautionary tales and proposed solutions. *npj Science of Learning*, *6*(1), 2. doi: 10.1038/s41539-020-00079-z

Hivert, V., Sidorenko, J., Rohart, F., Goddard, M. E., Yang, J., Wray, N. R., … Visscher, P. M. (2021). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *American Journal of Human Genetics*, *108*(5), 786–798. doi: 10.1016/j.ajhg.2021.02.014

Hong, H., Xu, L., Liu, J., Jones, W. D., Su, Z., Ning, B., … Shi, L. (2012). Technical reproducibility of genotyping SNP arrays used in genome-wide association studies. *PLoS ONE*, *7*(9), e44483. doi: 10.1371/journal.pone.0044483

Judd, N., Sauce, B., Wiedenhoeft, J., Tromp, J., Chaarani, B., Schliep, A., … Klingberg, T. (2020). Cognitive and brain development is independently influenced by socioeconomic status and polygenic scores for educational attainment. *Proceedings of the National Academy of Sciences of the USA*, *117*(22), 12411–12418. doi: 10.1073/pnas.2001228117

Kappelmann, N., Czamara, D., Rost, N., Moser, S., Schmoll, V., Trastulla, L., … Arloth, J. (2021). Polygenic risk for immuno-metabolic markers and specific depressive symptoms: A multi-sample network analysis study. *Brain Behavior and Immunity*, *95*, 256–268. doi: 10.1016/j.bbi.2021.03.024

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Karlsson Linnér, R., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., … Collaborators, C. (2021). Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nature Neuroscience*, *24*(10), 1367–1376. doi: 10.1038/s41593-021-00908-3

Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., von Stumm, S., Dale, P. S., … Plomin, R. (2018). Multi-polygenic score approach to trait prediction. *Molecular Psychiatry*, *23*(5), 1368–1374. doi: 10.1038/mp.2017.163

Martikainen, P., Korhonen, K., Jelenkovic, A., Lahtinen, H., Havulinna, A., Ripatti, S., … Silventoinen, K. (2021). Joint association between education and polygenic risk score for incident coronary heart disease events: A longitudinal population-based study of 26 203 men and women. *Journal of Epidemiology & Community Health*, *75*(7), 651–657. doi: 10.1136/jech-2020-214358.

Miguel, P. M., Pereira, L. O., Barth, B., de Mendonça Filho, E. J., Pokhvisneva, I., Nguyen, T. T. T., … Silveira, P. P. (2019). Prefrontal cortex dopamine transporter gene network moderates the effect of perinatal hypoxic-ischemic conditions on cognitive flexibility and brain gray matter density in children. *Biological Psychiatry*, *86*(8), 621–630. doi: 10.1016/j.biopsych.2019.03.983

Mitchell, B. L., Thorp, J. G., Wu, Y., Campos, A. I., Nyholt, D. R., Gordon, S. D., … Byrne, E. M. (2021). Polygenic risk scores derived from varying definitions of depression and risk of depression. *JAMA Psychiatry*, *78*(10), 1152–1160. doi: 10.1001/jamapsychiatry.2021.1988

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., … LifeLines Cohort, S. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449. doi: 10.1038/s41588-022-01016-z

Opel, N., Amare, A. T., Redlich, R., Repple, J., Kaehler, C., Grotegerd, D., … Dannlowski, U. (2020). Cortical surface area alterations shaped by genetic load for neuroticism. *Molecular Psychiatry*, *25*(12), 3422–3431. doi: 10.1038/s41380-018-0236-9

Østergaard, S. D., Trabjerg, B. B., Als, T. D., Climent, C. A., Privé, F., Vilhjálmsson, B. J., … Agerbo, E. (2020). Polygenic risk score, psychosocial environment and the risk of attention-deficit/hyperactivity disorder. *Translational Psychiatry*, *10*(1), 335. doi: 10.1038/s41398-020-01019-6

Plomin, R., & von Stumm, S. (2022). Polygenic scores: Prediction versus explanation. *Molecular Psychiatry*, *27*(1), 49–52. doi: 10.1038/s41380-021-01348-y

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. doi: 10.1038/ng.3285

Restrepo-Lozano, J. M., Pokhvisneva, I., Wang, Z., Patel, S., Meaney, M. J., Silveira, P. P., & Flores, C. (2022). Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. *Molecular Psychiatry*, *27*(6), 2742–2750. doi: 10.1038/s41380-022-01533-7

Ritchie, S. C., Lambert, S. A., Arnold, M., Teo, S. M., Lim, S., Scepanovic, P., … Inouye, M. (2021). Integrative analysis of the plasma proteome and polygenic risk of cardiometabolic diseases. *Nature Metabolism*, *3*(11), 1476–1483. doi: 10.1038/s42255-021-00478-5

von Stumm, S., Smith-Woolley, E., Ayorech, Z., McMillan, A., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Predicting educational achievement from genomic measures and socioeconomic status. *Developmental Science*, *23*(3), e12925. doi: 10.1111/desc.12925

Wainschtein, P., Jain, D., Zheng, Z., Aslibekyan, S., Becker, D., Bi, W., … Consortium, N. T.-O. f. P. M. (2022). Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nature Genetics*, *54*(3), 263–273. doi: 10.1038/s41588-021-00997-7

Warrier, V., Kwong, A. S. F., Luo, M., Dalvie, S., Croft, J., Sallis, H. M., … Cecil, C. A. M. (2021). Gene–environment correlations and causal effects of childhood maltreatment on physical and mental health: A genetically informed approach. *The Lancet. Psychiatry*, *8*(5), 373–386. doi: 10.1016/s2215-0366(20)30569-1

Winslow, E. B., & Shaw, D. S. (2007). Impact of neighborhood disadvantage on overt behavior problems during early childhood. *Aggressive Behavior*, *33*(3), 207–219. https://doi.org/10.1002/ab.20178

The failure of gene-centrism

Edward Archer[a] and Carl J. Lavie[b]

[a]Research and Development, EvolvingFX, LLC, Fort Wayne, IN, USA and [b]John Ochsner Heart & Vascular Institute, Ochsner Clinical School – The University of Queensland School of Medicine, New Orleans, LA, USA
archer1@evolvingfx.com
clavie@ochsner.org

doi:10.1017/S0140525X22002436, e209

Abstract

"Challenging the utility of polygenic scores for social science" is a compelling but limited critique. Phenotypic development is sensitive to both initial conditions *and all subsequent states* – from conception to senescence. Thus, gene-centric analyses are misleading (and often meaningless) because gene products are transformed, and their phenotypic 'effects' combined and attenuated with successive propagations from molecular and cellular contexts to organismal and social environments.

"Challenging the utility of polygenic scores for social science" is an erudite and compelling critique. Yet despite the analytic rigor, Burt failed to capture the fundamental flaws that render 'gene-centric' analyses misleading and polygenic scores (PGSs) meaningless outside of highly controlled environments.

The conflation of "inherited" with "genetic"

The functional unit in biology, biological inheritance, and phenotypic development is the cell – not DNA molecules. To be precise, humans develop from a single inherited cell – in which the genome is but one of many components. Moreover, because each cell's idiosyncratic nature and spatiotemporal context determine gene expression, the genome is merely an "*organ*" (McClintock, 1984, p. 800) or "*tool*" of the cell (Archer, 2015a, 2015b, 2015c). This fact renders the distinction between non-genetic (cellular) inheritance and the two forms of genetic inheritance (nuclear and mitochondrial) critical to analyses of phenotypic development. Nevertheless, gene-centric analyses ignore the distinction between cellular, nuclear, and mitochondrial inheritance, and the fact that intra- and extra-cellular environments determine both gene products and their 'effects.'

For example, the fundamental difference between monozygotic (identical) and dizygotic (fraternal) twins is inherent in the nomenclature – identical twins develop from a single cell (a fertilized egg) with a single placenta (usually); whereas fraternal twins develop from two different cells (two fertilized eggs) with two different placentas (always). Thus, fraternal twins differ in cellular and genetic inheritance, and their intrauterine environments, whereas identical twins do not.

Therefore, the greater phenotypic variability of fraternal twins is due to differences in gene expression engendered by *different cells* (eggs) acting in concert with inter-twin differences in both genotype and prenatal environments. Yet despite the extreme variability in the developmental competence (oocyte quality; e.g., mitochondrial content) of every female's population of eggs (Santos, El Shourbagy, & St John, 2006; Wang & Moley, 2010; Zhang et al., 2020; Zhou et al., 2020) and the irreversible impact of the intrauterine environment on development (Archer, 2015a, 2015b, 2015c, 2015d; Archer & Lavie, 2022; Archer, Lavie, Dobersek, & Hill, 2023), the functional distinction between cellular, genetic, and environmental (*in utero*) inheritance is absent in 'twin-studies,' heritability statistics, and polygenic scores (PGSs).

Thus, gene-centrism obscures the totality of biologic inheritance by conflating "genetic" with "inherited" – and therefore, the complexity of physiologic, psychological, and social phenotypic development remains unexamined.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Nonlinearity and development

Organismal development is a complex process that extends far beyond the linear amino acid sequence determined by the genetic code. For example, there are cellular, organismal, and environmental processes that lead to "one-to-many," "many-to-one," and "many-to-many" genotype–protein and genotype–phenotype relations. These processes include maternal and grandmaternal effects, phenotypic accommodation, alternative splicing, RNA editing, chimeric transcripts, protein multifunctionality, epistatic variance, the metabolic regulation of transcription, and post-translational modifications (Archer, Lavie, & Hill, 2018).

Therefore, because the genome does not have linear, predictive, or clear causal relations at the molecular level (e.g., protein species and function [Smith et al., 2021]), it is illogical to posit that it has these relations with physiologic and psychosocial phenotypes. Yet without these relations heritability statistics and PGSs are meaningless numbers.

Causality and non-genetic inheritance

Phenotypic trajectories are sensitive to initial conditions *and all subsequent states* – from conception to senescence. At each stage of development, extra-cellular environments alter intracellular environments – which then alter gene expression in a recursive process. Thus, because humans inherit molecules, cells, and their biologic, physical, and social environments from their parents, no single level of analysis (e.g., molecular, cellular, organismal, geographic, or societal) can be considered 'causal' unless the phenotypic changes at 'lower' levels of biologic organization can be shown to persist at 'higher' levels.

For example, the "egg" (primary oocyte) from which a human develops was initially created in the mother when she was a fetus developing in the grandmother's uterus. In other words, every "egg" that a female has was created prior to her birth. Thus, the physical and social environments in which a grandmother is immersed alter the intrauterine environment in which her offspring and the eggs of her female offspring develop. As such, physiologic, physical, and social environments impact the phenotypic development of at least three generations – the grandmother, her children, and her children's children – independent of the matrilineal genome.

These non-genetic processes of inheritance and evolution are known as "maternal and grandmaternal effects" and are well-established across species (Bateson et al., 2004; Gluckman, Hanson, Cooper, & Thornburg, 2008; Maestripieri & Mateo, 2009). For example, stunting and pediatric obesity *are caused by* adverse intrauterine environments – independent of genotype (Archer, 2015a, 2015b, 2015c, 2015d; Archer et al., 2023; Archer & Lavie, 2022; Archer & McDonald, 2017; Archer et al., 2018). In other words, starve any pregnant mammal and she will abort, or bear stunted offspring – independent of her genome.

Maternal effects and social outcomes

Importantly, disparities in egg (oocyte) quality and intrauterine environments can be caused by exposure to adverse physical and social environments (McQueen, Schufreider, Lee, Feinberg, & Uhler, 2015; Navot et al., 1991; Zhou et al., 2020). For example, low educational attainment, racism, sexism, and spousal abuse often lead to unremitting stress, poor nutrition, and alcohol, tobacco, or drug use that damage egg quality and the prenatal environments in which eggs develop. Thus, the environments and behaviors of past generations irreversibly alter the physiologic and behavioral phenotypes of current and future generations – independent of genotype.

Yet because the anatomic, physiologic, and psychological effects of prenatal insults are present at birth (inherited), persist into adulthood, and affect multiple generations, they are also inextricably linked to the matrilineal genome. Thus, gene-centric analyses that ignore maternal and grandmaternal effects will be misleading *because of* strong but demonstrably specious correlations between genotype and phenotype.

Given these facts, disparities in IQ, educational attainment, obesity, diabetes, physical activity, poverty, and criminal behavior in today's children are *caused by* the biologic, physical, and social environments in which their grandmothers and mothers were immersed from conception to senescence – independent of genomes and current environments. Thus, the adverse environments (and public policies) of yesteryear – not DNA – are causing disparities in health, wealth, and happiness today, and will continue to do so tomorrow.

Summary and conclusion

A great deal of biology – both established and undiscovered – links cellular and genetic inheritance (and the expression of that inheritance) with phenotypic development. Thus, estimates of genetic heritability and PGSs are often meaningless statistical abstractions derived from attempts to impose artificial dichotomies (nature vs. nurture and genes vs. environment) on demonstrably non-dichotomous developmental processes (Archer et al., 2018).

In closing, we agree with Burt's compelling critique and argue that an understanding of the etiology of physiologic, psychological, and social phenotypes "*will not be found in the genome*" (Archer, 2015a).

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Archer, E. (2015a). The childhood obesity epidemic as a result of nongenetic evolution: The maternal resources hypothesis. *Mayo Clinic Proceedings*, *90*(1), 77–92. https://doi.org/10.1016/j.mayocp.2014.08.006

Archer, E. (2015b). In reply – Maternal, paternal, and societal efforts are needed to "cure" childhood obesity. *Mayo Clinic Proceedings*, *90*(4), 555–557. https://doi.org/10.1016/j.mayocp.2015.01.020

Archer, E. (2015c). In reply – Epigenetics and childhood obesity. *Mayo Clinic Proceedings*, *90*(5), 693–695.

Archer, E. (2015d). The mother of all problems. *New Scientist*, *225*(3010), 32–33. https://doi.org/10.1016/S0262-4079(15)60404-3

Archer, E., & Lavie, C. J. (2022). Obesity subtyping: The etiology, prevention, and management of acquired versus inherited obese phenotypes. *Nutrients*, *14*(11), 2286. https://www.mdpi.com/2072-6643/14/11/2286

Archer, E., & Lavie, C. J. (2022). Obesity subtyping: The etiology, prevention, and management of acquired versus inherited obese phenotypes. *Nutrients*, *14*(11), 2286. http://dx.doi.org/10.3390/nu14112286

Archer, E., Lavie, C. J., Dobersek, U., & Hill, J. O. (2023). Metabolic inheritance and the competition for calories between mother and fetus. *Metabolites*, *13*(4), 545. http://dx.doi.org/10.3390/metabo13040545

Archer, E., Lavie, C. J., & Hill, J. O. (2018). The contributions of "diet", "genes", and physical activity to the etiology of obesity: Contrary evidence and consilience. *Progress in Cardiovascular Diseases*, *61*(2), 89–102. https://doi.org/10.1016/j.pcad.2018.06.002

Archer, E., & McDonald, S. M. (2017). The maternal resources hypothesis and childhood obesity. In M. S. Patel & J. S. Nielsen (Eds.), *Fetal and early postnatal programming and its influence on adult health* (pp. 17–32). CRC Press.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R. A., … Sultan, S. E. (2004). Developmental plasticity and human health. *Nature*, *430*(6998), 419–421. https://doi.org/10.1038/nature02725

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *New England Journal of Medicine*, *359*(1), 61–73. https://doi.org/10.1056/NEJMra0708473

Maestripieri, D., & Mateo, J. M. (Eds.). (2009). *Maternal effects in mammals*. University of Chicago Press.

McClintock, B. (1984). The significance of responses of the genome to challenge. *Science*, *226*(4676), 792–801. http://www.jstor.org/stable/pdfplus/1693251.pdf?acceptTC=true

McQueen, D. B., Schufreider, A., Lee, S. M., Feinberg, E. C., & Uhler, M. L. (2015). Racial disparities in in vitro fertilization outcomes. *Fertility and Sterility*, *104*(2), 398–402.e391. https://doi.org/10.1016/j.fertnstert.2015.05.012

Navot, D., Bergh, R. A., Williams, M. A., Garrisi, G. J., Guzman, I., Sandler, B., & Grunfeld, L. (1991). Poor oocyte quality rather than implantation failure as a cause of age-related decline in female fertility. *The Lancet*, *337*(8754), 1375–1377. https://doi.org/10.1016/0140-6736(91)93060-M

Santos, T. A., El Shourbagy, S., & St John, J. C. (2006). Mitochondrial content reflects oocyte variability and fertilization outcome. *Fertility and Sterility*, *85*(3), 584–591. https://doi.org/10.1016/j.fertnstert.2005.09.017

Smith, L. M., Agar, J. N., Chamot-Rooke, J., Danis, P. O., Ge, Y., Loo, J. A., … Kelleher, N. L. (2021). The human proteoform project: Defining the human proteome. *Science Advances*, *7*(46), eabk0734. https://doi.org/10.1126/sciadv.abk0734

Wang, Q., & Moley, K. H. (2010). Maternal diabetes and oocyte quality. *Mitochondrion*, *10*(5), 403–410. https://doi.org/10.1016/j.mito.2010.03.002

Zhang, J.-J., Liu, X., Chen, L., Zhang, S., Zhang, X., Hao, C., & Miao, Y.-L. (2020). Advanced maternal age alters expression of maternal effect genes that are essential for human oocyte quality. *Aging* (*Albany, NY*), *12*(4), 3950–3961. https://doi.org/10.18632/aging.102864

Zhou, X., McQueen, D. B., Schufreider, A., Lee, S. M., Uhler, M. L., & Feinberg, E. C. (2020). Black recipients of oocyte donation experience lower live birth rates compared with white recipients. *Reproductive BioMedicine Online*, *40*(5), 668–673. https://doi.org/10.1016/j.rbmo.2020.01.008

Complex interactions confound any unitary approach to social phenomena, not just biological ones

Darren Burke

School of Psychological Science, University of Newcastle, Ourimbah, NSW, Australia
Darren.burke@newcastle.edu.au; https://darrenburke.wixsite.com/evocoglab-1

doi:10.1017/S0140525X22002369, e210

Abstract

Although Burt clearly explains how modern genomic techniques work, and describes their limitations, her conclusion that they are therefore not valuable additions to understanding social outcomes is unwarranted. Understanding the causes of complex social outcomes depends on understanding how social, individual, and genetic factors complexly interact with each other. None can be understood without reference to the others.

Burt does an outstanding job of clearly describing, at just the right level of necessary detail, how modern genomic techniques work, and clearly articulates the many limitations associated with drawing strong inferences about the role played by genes in determining any particular behavioral or social outcome. She also acknowledges that most researchers who use these techniques (even the "enthusiasts") are aware of these limitations, and typically do what they can to mitigate them. The crux of her argument is that these attempts are inadequate (maybe even unavoidably so), and so continued efforts in that direction are a waste of limited resources, and potentially dangerous, because they might be "obscuring social–structural and physical environmental influences and promoting the individualization of social problems" (target article, sect. 8, para. 4). These perceived potential dangers reveal a laudable (but perhaps misguided) political motivation for discouraging continued research into genetic influences, but Burt is explicit about the fact that she is not arguing from a political perspective, but a scientific one. From a scientific perspective, deliberately ignoring genetic influences on even the most complex and apparently purely "social" phenomena is a straightforward mistake.

One reason for rejecting Burt's proposal on scientific grounds is that it rules out the possibility of understanding the evolutionary origin of complex social phenomena, which, of course, depends on explicating the role genes might play in such effects. The reason this is important is because the evolutionary selected consequences of a gene's actions *are* its effects. Burt uses Jenks' (1972) classic thought experiment about discrimination based on hair color to bolster her argument for purely social factors creating what look like genetic influences on a social outcome. The thought experiment asks to imagine a world in which red-haired children are barred from school. Under such conditions, there would be a strong relationship between genes (those responsible for pheomelanin production; Valverde, Healy, Jackson, Rees, & Thody, 1995) and educational attainment, but Burt's argument is that this relationship would be spurious, because the genes themselves do not "code" for anything directly related to an individual-level propensity to acquire an education (e.g., intelligence, or perseverance, or carefulness). Although it is intended to illustrate the opposite point, this thought experiment can be used to exemplify the fundamental premise of evolution; that genes *only* have selective consequences *in* environments – this *interaction* underpins *all* selection. This example makes it obvious that the selective consequences of the gene need not be direct. Indeed, as genes only make proteins (or regulate the genes that make proteins), it *can* never be direct. From an evolutionary perspective, if school-level education increased reproductive success (increased "fitness" in evolutionary biology), for *whatever* reason, then the genes for red hair would be selected against, *because* they reduced educational opportunities. This *is* one of the gene's consequences in that environment, and so there is a real sense in which this is one of its *effects*, despite being indirect. The only reason this looks like a spurious association is because we know that the hypothetical educational barrier is artificially constructed. We rarely have that insight for most genes and their real effects, but understanding all evolution depends on understanding the effects of genes in environments. For example, imagine a world in which rather than red-heads being barred from school, they were considered more attractive, and so more easily attracted mating partners. Again, whether this preference was because of a sexually selected "genetic" preference (Endler & Basolo, 1998) or an entirely arbitrary socially determined preference (if such a thing could exist), the possession of the genes for making red hair would have evolutionary

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

consequences, and those consequences would *be* one of its effects.

A major theme of Burt's critique is that modern genomic techniques are guilty of neglecting the kind of developmental interactionism described above, and have a tendency to ignore the fact that such findings are therefore context- and population-specific. Although this criticism may be true of the most ardent enthusiasts, the vast majority of researchers in this field clearly demonstrate awareness of these issues by trying to factor them in (or out) in their analyses. The irony of this criticism is that Burt cites a host of social factors that are presented as causes of social outcomes without any direct evidence of their casual efficacy and without any acknowledgement that those factors can only have their effects via interacting with genes. It is true, as Burt points out, that there is no meaningful way we can identify social-context-independent genetic "potential" for a particular social outcome (like educational attainment), but it is equally true that there is no meaningful way to identify gene-independent social "potential" for any social outcome. No matter what socially determined education-relevant advantages (or disadvantages) a person has, that can only possibly translate into actual educational attainment via a long, complicated interaction between their genes and those social factors. Part of the problem here is that "social" variables are discussed as causes and outcomes, without a complete fleshing out of the individual-level causal mechanisms involved in linking them. It is *individuals* who attain certain levels of education, or have same-sex sex, or who play golf, to use some of Burt's examples, and so we need to understand how genes and environments dynamically interact to understand why individuals behave in certain ways in certain environments. Two obvious individual-level factors that influence many of these outcomes are personality and intelligence, and both of those are traits that are a consequence of complex gene–environment interactions. For any outcome we might be interested in understanding, *all* of these factors (social, individual, and genetic) complexly interact, and so there are logistical barriers to identifying "the role" played by any one factor, and logical reasons to avoid any such approach. We need good data on all of the factors involved, and so discouraging any approach is counterproductive.

Competing interest. None.

References

Endler, J. A., & Basolo, A. L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in Ecology and Evolution*, *13*, 415–420.

Valverde, P., Healy, E., Jackson, I., Rees, J. L., & Thody, A. J. (1995). Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genetics*, *11*(3), 328–330. doi: 10.1038/ng1195-328, PMID 7581459. S2CID 7980311.

Downward causation and vertical pleiotropy

Evan Charney

Samuel DuBois Cook Center on Social Equity, Duke University, Durham, NC, USA
echar1@me.com

doi:10.1017/S0140525X22002308, e211

Abstract

In discussing the relationship between genetically influenced differences and educational attainment (EA), Burt employs the concept of downward causation. I note the similarities between Burt's concept of downward causation and the sociogenomics concept of vertical pleiotropy and argue that her discussion of downward causation introduces an unnecessary normative component. The core problem concerns not the appropriateness of phenotypes that influence EA but mistaken assumptions about which phenotypes are being predicted.

Based on Burt's definition of downward causation, phenotypes A and B exhibit downward causation when risk alleles for phenotype A predict phenotype B. A's risk alleles predict B neither because B is a biological consequence of A (e.g., kidney disease can be a biological consequence of diabetes) nor because A and B share risk alleles (e.g., multiple risk loci are shared between autoimmune diseases). Instead, the causal connection between A and B is because of socio-environmental forces (norms, practices, institutions). In addition, as the example of educational attainment (EA) shows, B may be a wholly social construct without its own risk alleles.

Thus characterized, downward causation is equivalent to what sociogenomicists misleadingly refer to as "vertical pleiotropy" (van Rheenen, Peyrot, Schork, Lee, & Wray, 2019). Pleiotropy occurs when a single gene plays a causal role in two or more distinct phenotypes. Despite its name, vertical pleiotropy is not a form of pleiotropy at all because phenotypes A and B do not share any causal alleles. Hill and Davis, for example, treat genetic variants that ostensibly predict income as an example of vertical pleiotropy, noting that (2019, p. 19), "genetic variants do not act directly on income; instead, genetic variants are associated with partly heritable traits (such as intelligence, conscientiousness, health, etc.), which have their own complex gene-to-phenotype paths (including neural variables) and are ultimately associated with income." They also comment that any correlation between a given attribute (e.g., intelligence or health) and income is determined by social institutions: Income could just as well depend on service to the party, and different political policies could alter, if not eliminate, the degree of correlation between, for example, income and health.

As with income, most sociogenomicists would agree there are no alleles for EA *per se*. Rather, there are alleles for attributes that causally affect EA and it is the socially constructed nature of the educational system that determines what attributes of persons are relevant and rewarded (which may include features of persons that are socially valued, such as attractiveness and height, but not knowingly made criterion of EA). Perhaps in one society, obedience is valued more than critical thinking and rewarded accordingly. Under the assumption that cognitive performance is "strongly" genetically influenced and that socially it exerts a decisive causal influence on EA, sociogenomicists typically treat EA as a "proxy variable" for cognitive performance (Rietveld et al., 2014, p. 13791).

Burt implies that in addition to the properties mentioned above, downward causation is characterized by phenotype A being an *inappropriate* socially mediated cause of EA. All the examples she presents for phenotype A – ethnicity, skin color, attractiveness, height, weight – are examples in which most would agree that it is indeed inappropriate, if not a grave social injustice, that A has a causal effect on EA. Burt notes of such phenotypes, "In a GWAS, such alleles [alleles associated with skin pigmentation of African Americans, but also attractiveness, height, weight, etc.] would be

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

identified as causing differences in educational attainment, but these association signals would, of course, be artificial." However, the signals would be no less "artificial" if the alleles identified as causing differences in EA were associated with *intelligence*. Although many would consider this an "appropriate" cause of differences in EA, as noted above, to the extent that it is a cause is because of contingent social and institutional practices and norms.

Normative objections in this context makes one vulnerable to the charge (common enough) that one's objection is not scientific. Such an invocation is unnecessary because Burt herself has already convincingly demonstrated the problem with downward causation in the context of EA (and most other social attributes such as income). The problem lies with the assumption that *EA is a proxy variable for intelligence*, that is, that in measuring EA, sociogenomicists are measuring intelligence. The error is scientific, not normative, to the extent that this assumption is wrong.

First, as Burt shows, to whatever extent intelligence has genetic influences, the realization of intelligence as a phenotype is so intertwined with so many socio-environmental variables that it is impossible to separate "the genetic influence on intelligence" as some pre-existing, isolated, potential force. Moreover, one of these influences on intelligence may well be education itself. In place of the assumed unidirectional causal pathway from intelligence to EA, EA itself may influence intelligence, resulting in reciprocal or bidirectional causation (Hegelund et al., 2020).

Second, as Burt also shows, there are strong reasons to believe that polygenic EA scores predict not intelligence but *ancestry*. Population stratification itself is an example of downward causation/vertical pleiotropy. Genetic ancestry bears a socially determined association with any number of social attributes, EA and income being two noteworthy examples. In the latest in a long series of studies showing the enduring impact of population stratification on genome-wide association studies (GWASs) of complex traits, the authors note that "controlling for geographic regions significantly decreased the heritability for socioeconomic status (SES)-related traits, most strongly for educational attainment and income" (Abdellaoui, Dolan, Verweiji, & Nivard, 2022).

A final word concerning sociogenomicists' repeated assertion that in addition to the heritability of intelligence (and whatever other attributes are considered to have an association with EA), *EA is itself heritable* (the same is said of income). How can EA be said to be heritable if there are no genetic variants that act directly on it? One might object that nothing in the concept of heritability requires that a trait deemed heritable be influenced by the transmission of parental risk alleles *for that trait*. It is sufficient that heritable trait A stands in a (socially mediated) causal relationship to trait B. However, if we accept this, we would have no grounds to claim, to use an example of population stratification cited by Burt, that chopstick use is not heritable. Rather, we could say that it is an example of vertical pleiotropy. Being of East Asian descent (phenotype A) is a heritable attribute, and because of social practices (norms, conventions) it is causally associated with chopstick use (phenotype B).

Financial support. The author has received no funding for this article.

Competing interest. None.

References

Abdellaoui, A., Dolan, C. V., Verweiji, K. J. H., & Nivard, M. G. (2022). Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature Genetics*, *54*, 1345–1354. https://doi.org/10.1038/s41588-022-01158-0

Hegelund, E. R., Grønkjær, M., Osler, M., Dammeyer, J., Flensborg-Madsen, T., & Mortensen, E. L. (2020). The influence of educational attainment on intelligence. *Intelligence*, *78*, 101419. https://doi.org/10.1016/j.intell.2019.101419

Hill, W. D., Neil, M. D., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., … Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*(1), 5741. https://doi.org/10.1038/s41467-019-13585-5

Rietveld, C. A., Esko, T., Davies, G., Pers, T. H., Turley, P., & Benyamin, B. (2014). Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(38), 13790–13794. https://doi.org/10.1073/pnas.1404623111

van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., & Wray, N. R. (2019). Genetic correlations of polygenic disease traits: From theory to practice. *Nature Reviews Genetics*, *20*(10), 567–581. https://doi.org/10.1038/s41576-019-0137-z

Social scientists would do well to steer clear of polygenic scores

David Curtis[a,b]

[a]UCL Genetics Institute, University College London, London, UK and [b]Centre for Psychiatry, Queen Mary University of London, London, UK
d.curtis@ucl.ac.uk; http://www.davecurtis.net/dcurtis.html

doi:10.1017/S0140525X22002448, e212

Abstract

The problems with polygenic scores (PGSs) have been understated. The fact that they are ancestry-specific means that biases related to sociodemographic factors would be impossible to avoid. Additionally, the requirement to obtain DNA would have profound impacts on study design and required resources, as well as likely introducing recruitment bias. PGSs are unhelpful for social science research.

Burt does an excellent job of debunking some of the hype attaching to sociogenomics and the field of polygenic scores (PGSs) in general. Although she concludes that PGSs may be not very useful for social science, in fact there are good reasons to regard them as perhaps being worse than useless.

Why should social scientists feel quite comfortable not incorporating PGSs into their research? There is no doubt that genetic factors can have substantial effects on relevant outcomes. For example, countless variants in DNA sequence have been identified which lead to profound intellectual disability, effectively reducing educational attainment to zero (Ilyas, Mir, Efthymiou, & Houlden, 2020). Likewise, it is not up for debate that the effect of some genetic variation will be moderated by environment. Genetic factors increasing athletic ability will be expected to be associated with increased educational attainment if colleges recruit students on sports scholarships and less so if admission is based only on intellectual capability. So the issue is not that genetic factors do not impact outcomes of interest but rather, as Burt explains, that PGSs are so poor at capturing the genetic variation which is biologically relevant while at the same time being profoundly influenced by exactly the kind of confounders social scientists do not want contaminating their research, such as race, socioeconomic status, and parental characteristics.

Although Burt does touch on many of the relevant issues, I would argue that the situation is even more problematic than

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

she presents it to be. In my view, what she refers to as population stratification produces effects of such magnitude and malignancy as to render the proposal to routinely incorporate PGSs as covariates in social science research a complete non-starter. There are two related phenomena. One is that the absolute magnitude of PGSs varies with ancestry and the other is that the strength of the association between a PGS and the trait it is supposed to predict also varies with ancestry (Martin et al., 2019). These are not small effects. The PGS for schizophrenia is much more strongly associated with ancestry than it is with schizophrenia (Curtis, 2018). Researchers working with PGSs now routinely use ancestry-specific PGSs produced by carrying out genome-wide association studies (GWASs) in relevant cohorts. A PGS for white Europeans will need to be derived from a GWAS of an exclusively white European cohort; a PGS for Asians will be derived from a GWAS of an exclusively Asian cohort (Ho et al., 2022). And so on, except that because Africans have more genetic diversity than other populations a PGS derived from a GWAS of an African cohort will always perform less well than its counterparts for other ancestries.

Given these now well-recognised properties of PGSs it is truly challenging to see how one could consider incorporating a PGS as a covariate in a social science research project. The value of a subject's PGS would be profoundly influenced by their ancestry. If one went down the route of attempting to use an ancestry-specific PGS then a prerequisite would be that a GWAS of the trait in question should have been performed on every relevant ancestry group. Knowing which one to use would require determining the ancestry of each subject. For subjects of mixed ancestry, an attempt would need to be made to combine PGSs (Marnetto et al., 2020). For subjects with African ancestry the PGS would capture less of the genetic risk than for other subjects. Thus, the PGS represents a variable which not only performs badly in terms of measuring genetic risk but also performs more badly for some subjects than others. Such an obvious source of systematic bias would make it difficult or impossible to draw useful conclusions from studies which incorporated it.

There is another way in which Burt's treatment is too kind to PGSs. She has not presented a full account of the difficulties of obtaining them for participants in a social science research project. Once one has obtained single-nucleotide polymorphism (SNP) genotypes then producing a PGS is a trivial exercise. But obtaining SNP genotypes cannot be done by having the subject fill out a questionnaire or go through a structured interview – they have to actually donate a DNA sample and it has to be processed by a laboratory. Incorporating a PGS into social science research involves adding a whole new biological dimension with a very substantial impact on the overall shape of the project. It also requires that subjects voluntarily provide a DNA sample. Although some may be happy to do this, it is unarguably the case that a DNA sample represents a large quantity of personal information which is potentially sensitive in a number of ways (Alsaffar, Hasan, McStay, & Sedky, 2022). An individual's genetic profile provides at least some information about their risk of a large number of health conditions. It could potentially be of use to police and security forces who might seek to identify the perpetrator of a crime, or at least one of their relatives. Although safeguards may be in place which attempt to prevent the misuse of genetic data some individuals may feel reluctant to provide DNA for reasons which are not wholly irrational. Of especial concern is that one might well expect that factors influencing an individual's enthusiasm for donating DNA would include a number of factors which might be of interest to social scientists, such as education, race, health, substance misuse, and criminality. Thus, introducing DNA sampling as a routine aspect of social science research seems certain to introduce systematic bias into recruitment. And as far as research involving children is concerned, I would argue that the privacy concerns about possible misuse of genetic data would mean that it could not be ethical to obtain DNA even if their parents consented.

The inclusion of PGSs into social science research is impractical and highly likely to introduce bias. For these reasons and others, I believe that PGSs have a negative utility.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Alsaffar, M. M., Hasan, M., McStay, G. P., & Sedky, M. (2022). Digital DNA lifecycle security and privacy: An overview. *Briefings in Bioinformatics*, *23*(2), 1–16. https://doi.org/10.1093/BIB/BBAB607

Curtis, D. (2018). Polygenic risk score for schizophrenia is more strongly associated with ancestry than with schizophrenia. *Psychiatric Genetics*, *28*(5), 85–89. https://doi.org/10.1097/YPG.0000000000000206

Ho, W. K., Tai, M. C., Dennis, J., Shu, X., Li, J., Ho, P. J., … Teo, S. H. (2022). Polygenic risk scores for prediction of breast cancer risk in Asian populations. *Genetics in Medicine*, *24*(3), 586–600. https://doi.org/10.1016/J.GIM.2021.11.008

Ilyas, M., Mir, A., Efthymiou, S., & Houlden, H. (2020). The genetics of intellectual disability: Advancing technology and gene editing. *F1000Research*, *9*. https://doi.org/10.12688/F1000RESEARCH.16315.1

Marnetto, D., Pärna, K., Läll, K., Molinaro, L., Montinaro, F., Haller, T., … Pagani, L. (2020). Ancestry deconvolution and partial polygenic score can improve susceptibility predictions in recently admixed individuals. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-15464-w

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*(4), 584–591. https://doi.org/10.1038/s41588-019-0379-x

Often wrong, sometimes useful: Including polygenic scores in social science research

Jason Fletcher

Center for Demography of Health and Aging, La Follette School of Public Affairs, University of Wisconsin-Madison, Madison, WI 53706, USA
jason.fletcher@wisc.edu

doi:10.1017/S0140525X22002461, e213

Abstract

This commentary seeks to briefly outline a clear-eyed middle ground between Burt's claims that the inclusion of polygenic scores (PGSs) is essentially useless for social science and proponents' vast overstatements and over-interpretations of these scores. Current practice of including PGSs in social science is often wrong but sometimes useful.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Burt's goal is to challenge arguments about the value of including genetic measurements in social science research. The author focuses on a subset of "genetics" – the use of polygenic scores (PGSs) and lists four key "limitations of PGS that undermine their utility for social science." I'll summarize these as (1) they are not "purely" genetics and are thus confounded (2) the causal mechanism is unclear; in cases where the mechanism is environmental, this is a labeled "downward causation" which is said to produce "artificial" genetic signals (3) they are incomplete measures of genetic variation and (4) their interpretation is context-dependent. On their face, these four limitations would seemingly apply to, essentially, all variables used in social science research – and this is a key double-edged sword of exploring the use of "genetics" in social science: To on the one hand treat them as special and on the other treat them as "regular" variables. Proponents want them to be treated as regular when evaluating their general use and special when interpreting their effects and opponents want the opposite. Like other commenters, Burt's arguments are too unfocused and often imprecise, in my view; focusing on the dissonance of opponent's treatment of these variables without acknowledging the dissonance in her own arguments. The arguments lack the specificity that is needed and conflates issues of *interpretation* of the PGS in an empirical application with the net-scientific-value of including the PGS at all. Instead, I believe the two key features of using PGSs are its utility in the specific application and a need to under-, rather than over-, interpret the PGS as "genetic" effects[1] at all. Like models, PGS inclusion can be wrong but useful. Unfortunately, many proponents want to leverage the "wrong" to strengthen arguments of the importance of genetics more broadly and put less emphasis on the "useful."

Burt is absolutely correct that the ambiguous nature of a PGS's interpretation has led far too many investigators to over-interpret and narrowly label a PGS as "genetic," often to elevate the perceived importance of "genetics" in contributing to social science outcomes.[2] For example, many investigators aspire to specifically distinguish a "genetic" effect from an effect stemming from a broader "family background" source. At present, I believe this effort is a fool's errand and research that attempts such a separation should be understood as over-stepping and over interpreting and largely dismissed as such (Fletcher, Wu, Li, & Lu, 2021).

However, let's return to some purported uses of PGSs that may shed light rather than only muddying the waters. In many investigations of whether an environmental exposure affects an outcome, researchers are worried that some "third factor" might cause both the environment and outcomes. In many such analyses, a standard and reasonable question is whether genetics and/or family background is the "third factor." A very standard approach to partially address this specific concern is to compare siblings (i.e., hold constant shared family background and shared genetics). This approach is useful but imperfect (e.g., Boardman & Fletcher, 2015). In some circumstances it can provide useful, directional evidence of the importance of this particular third factor source. As an alternative – in a situation without sibling data, for example – researchers could instead control for PGS, perhaps in conjunction with a more formal sensitivity analysis of the original results (Oster, 2019). If the researchers do not attempt to interpret the effect of this third factor on the outcome (i.e., "genetic effects"), which follows standard practice in interpretations of an included potential third factor, then I believe the inclusion of PGS can be quite useful in standard social science analysis. A second use is, essentially, imputation of variables the researcher does not have in the data. Again, the focus is not on interpreting the "effects" of the PGS, but using them as signals for where to collect social science data in the future. Burt describes cases where a researcher has data on both a trait of interest and a relevant PGS and chooses to use the PGS (e.g., school grades) in a downstream analysis of, say, predicting high school graduation. But what about the case where school grades are not measured? Or when school grades are only measured post-treatment (e.g., for an early-life intervention)? In these cases a PGS – and even better, many different PGSs – could be used for hypothesis generation for future analysis. For example, an early (or *in utero*) intervention that is shown to interact differentially with PGS for cognition, PGS for ADHD, PGS for risk tolerance, and so on in predicting high school graduation could be both wrong (if we try to interpret the "genetic effects" directly) but useful (if we do not).

Overall, Burt's paper summarizes a useful set of issues around the inclusion of PGS measures in social science research. I believe the issues raised are mostly correct when focusing on the broad misinterpretation of PGSs as representing "genetic effects" in the emerging literature. However, I believe these misinterpretations can be challenged and corrected directly without the need to abandon the inclusion of PGSs is a limited and focused role in social science research.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

Notes

1. As Boardman and Fletcher (2021) state "…even if each of the genetic variant effects that are added together were causal effects, the resulting summary measure [(i.e. PGS)] would not have a clear interpretation. Many researchers have used vague terms, such as *genetic endowment*, *genetic risk*, or *genetic predisposition*, in labeling these constructs… the fact that many of the genetic variant effects are not causal further challenges the interpretation – so much so that it is not clear that they can be called 'genetic' effects at all…"
2. As described in more detail in Fletcher (2022a, 2022b), Harden (2021) is a particularly poignant case of over-stating current knowledge and methods. Likewise, the new chapter by Madole and Harden (2023) overstates and oversells.

References

Boardman, J. D., & Fletcher, J. M. (2015). To cause or not to cause? That is the question, but identical twins might not have all of the answers. *Social Science and Medicine*, *127*, 198–200.

Boardman, J. D., & Fletcher, J. M. (2021). Evaluating the continued integration of genetics into medical sociology. *Journal of Health and Social Behavior*, *62*(3), 404–418.

Fletcher, J. (2022a). Backdoor to a dead end: A review essay: The genetic lottery: Why DNA matters for social equality by Kathryn Paige Harden. *Population and Development Review*, *48*(1), 253–258.

Fletcher, J. (2022b). Lost in translation: Importing causal designs into behavioral genetics may be a dead end. https://osf.io/preprints/socarxiv/zyvkr/

Fletcher, J., Wu, Y., Li, T., & Lu, Q. (2021). Interpreting polygenic score effects in sibling analysis. *BioRxiv*.

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Madole, J., & Harden, K. P. (2023). Building causal knowledge in behavior genetics. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X22000681

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Polygenic scores, and the genome-wide association studies they derive from, will have difficulty identifying genes that predispose one to develop a social behavioral trait

Edward Fox

Department of Psychological Science, Purdue University, West Lafayette, IN, USA
au_gc@psych.purdue.edu
https://www.purdue.edu/hhs/psy/directory/faculty/Fox_Edward.html

doi:10.1017/S0140525X22002370, e214

Abstract

Polygenic scores (PGSs) have several limitations. They are confounded with environmental effects on behavior and cannot be used to study how mutations affect brain function and behavior. For this, mutations with large effects, which often arise in only one geographical population are needed. Genome-wide association studies (GWASs), commonly used for identifying mutations, have difficulty detecting these mutations. A strategy that overcomes this challenge is discussed.

Proponents of sociogenomics argue polygenic scores (PGSs) should be incorporated into social science research. PGSs are derived from many variants (up to thousands) of very small effect size (common variants) that are associated with measures of a social behavior trait as determined by genome-wide association studies (GWASs; Pain et al., 2021). As Burt described, although PGSs are suggested to be measures of genetic influence or propensity for complex traits, several factors make it difficult or impossible to distinguish genetic and environmental effects on such traits. Thus, PGSs are unlikely to be strictly genetic predictors of the propensity to exhibit a trait. Moreover, PGSs do not typically identify alleles or variants responsible for a phenotype (Astle et al., 2016). This is unfortunate because identifying a deleterious mutation would permit its biological activity to be studied. The information obtained could provide the knowledge needed to repair or counteract the deleterious effects of the mutation. This limitation of PGSs is exemplary of a broader issue as GWASs have identified thousands of strong associations with complex diseases and traits, but in very few instances has the actual risk variant been identified (Chorley et al., 2008), or have they been successfully translated into clinical use (Bomba, Walter, & Soranzo, 2017). Identifying causal common variants in GWASs has been difficult because they usually map to regulatory regions (Astle et al., 2016), where they influence gene expression, including processes involved in execution of gene expression such as splicing (Lalonde et al., 2011).

In contrast to common variants of small effect size, rare variants that have large effects on phenotypes have been identified. These variants are often associated with the protein-coding portion of a gene. As proteins are important for structural and physiological functions of cells, mutations that affect them can produce these large effects. An explanation for the rarity of these variants based on evolutionary theory proposes that the detrimental effect of disease on fitness results in selection against variants that promote disease (Gibson, 2012).

Rare variants are often identified by quantitative trait locus (QTL) analysis, which looks for correlations between variants and measures of continuous phenotypic traits (Bloom et al., 2019). The goal is to uncover the locations in the genome important for these traits. A variation of this analysis that has identified rare variants of large effect size used individuals that displayed the trait of interest and individuals that did not display it from multiple generations of families or isolated populations. Rare variants might be found at higher frequencies in isolated populations because of previous bottleneck events, genetic drift or adaptation, and selection (Moltke et al., 2014). This increases the power to detect associations between rare variants and phenotypes (Colonna et al., 2013). In these studies that sample from families or isolated populations, variants that are closely linked to the mutation or causative allele are present in individuals that exhibit the trait at higher frequency than in individuals that do not display the trait. The locations of these variants indicate the chromosome region likely to contain the mutation. Positional cloning within this region can be used to identify the mutated gene and then comparison of this gene's DNA sequence in subjects with and without the trait can identify the causative mutation. Even though this mutation might only be present in a family or isolated population, the ability to study how any mutation alters the brain to influence a complex behavioral trait would be a breakthrough. An example of a study with success using this strategy focused on Canadian families of Celtic descent with multiple relatives in up to three generations diagnosed for schizophrenia (Brzustowicz, Hodgkinson, Chow, Honer, & Bassett, 2000). A highly significant association between schizophrenia and a locus on chromosome 1q21–q22 was found. Then additional variants within this region were used to pinpoint the nitric oxide synthase 1 adaptor gene (Brzustowicz et al., 2004). This gene is overexpressed in the frontal cortex of people with schizophrenia, and it is involved in synaptic function and cortical neuron development, effects that could contribute to schizophrenia (Carrel et al., 2015; Hernandez et al., 2016).

In contrast, GWASs, and thus PGSs, do not typically detect QTLs or rare variants of large effect size because these variants are rare in the total population sampled by GWASs. The power to detect a variant of any effect size decreases with the frequency of the variant because fewer individuals in the sample carry a less-frequent variant (Zuk et al., 2014). Put another way, because GWASs calculate the average effects of alleles across thousands of individuals, they cannot capture heterogeneity of effect sizes at the family level (Gibson, 2012).

Can approaches that detect rare variants be useful for sociogenomics? It could be argued that some measures of interest in sociogenomics, for example, level of educational attainment, could not be accounted for by one or a few rare variants. However, the contrast between what GWASs and PGSs identify best (common variants of small effect size) versus what QTL and related approaches identify best (rare variants of large effect size) suggests QTL and related approaches could have significant relevance for sociogenomics. As discussed above, by studying the right population it may be possible to identify associations of a complex behavioral trait with rare variants of large effect and ultimately identify one or more causative alleles. Social behaviors are complex and depend on multiple interacting neural systems as illustrated in a recent review on neural encoding

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

of social valence (Padilla-Coreano, Tye, & Zelikowsky, 2022). Social attributes, social memory, social rank, and social isolation were proposed to influence valence assignment to social stimuli, which in turn influences social interactions. Also, the separate neural circuits that control each of these influences were described, noting some overlap of these circuits. Interestingly, they suggest that across psychiatric disorders, brain regions that contribute to encoding of valence and social functions exhibit abnormal activity during emotional processing (e.g., Laviolette, 2007). Thus, if a mutation disrupts one or more of the neural systems that influences valence assignment, this might lead to abnormal social interactions and a search might identify causal variants, including rare ones of large effect size.

Financial support. This research received no special grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., … Soranzo, N. (2016). The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell*, *167*(5), 1415–1429, e1419. doi: 10.1016/j.cell.2016.10.042

Bloom, J. S., Boocock, J., Treusch, S., Sadhu, M. J., Day, L., Oates-Barker, H., & Kruglyak, L. (2019). Rare variants contribute disproportionately to quantitative trait variation in yeast. *eLife*, *8*. https://doi.org/10.7554/eLife.49212

Bomba, L., Walter, K., & Soranzo, N. (2017). The impact of rare and low-frequency genetic variants in common disease. *Genome Biology*, *18*(1), 77. doi: 10.1186/s13059-017-1212-4

Brzustowicz, L. M., Hodgkinson, K. A., Chow, E. W., Honer, W. G., & Bassett, A. S. (2000). Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21–q22. *Science (New York, N.Y.)*, *288*(5466), 678–682. doi: 10.1126/science.288.5466.678

Brzustowicz, L. M., Simone, J., Mohseni, P., Hayter, J. E., Hodgkinson, K. A., Chow, E. W., & Bassett, A. S. (2004). Linkage disequilibrium mapping of schizophrenia susceptibility to the CAPON region of chromosome 1q22. *American Journal of Human Genetics*, *74*(5), 1057–1063. doi: 10.1086/420774

Carrel, D., Hernandez, K., Kwon, M., Mau, C., Trivedi, M. P., Brzustowicz, L. M., & Firestein, B. L. (2015). Nitric oxide synthase 1 adaptor protein, a protein implicated in schizophrenia, controls radial migration of cortical neurons. *Biological Psychiatry*, *77*(11), 969–978. doi: 10.1016/j.biopsych.2014.10.016

Chorley, B. N., Wang, X., Campbell, M. R., Pittman, G. S., Noureddine, M. A., & Bell, D. A. (2008). Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: Current and developing technologies. *Mutation Research*, *659*(1–2), 147–157. doi: 10.1016/j.mrrev.2008.05.001

Colonna, V., Pistis, G., Bomba, L., Mona, S., Matullo, G., Boano, R., … Toniolo, D. (2013). Small effective population size and genetic homogeneity in the Val Borbera isolate. *European Journal of Human Genetics*, *21*(1), 89–94. doi: 10.1038/ejhg.2012.113

Gibson, G. (2012). Rare and common variants: Twenty arguments. *Nature Review Genetics*, *13*(2), 135–145. doi: 10.1038/nrg3118

Hernandez, K., Swiatkowski, P., Patel, M. V., Liang, C., Dudzinski, N. R., Brzustowicz, L. M., & Firestein, B. L. (2016). Overexpression of isoforms of nitric oxide synthase 1 adaptor protein, encoded by a risk gene for schizophrenia, alters actin dynamics and synaptic function. *Frontiers in Cellular Neuroscience*, *10*, 6. doi: 10.3389/fncel.2016.00006

Lalonde, E., Ha, K. C., Wang, Z., Bemmo, A., Kleinman, C. L., Kwan, T., … Majewski, J. (2011). RNA sequencing reveals the role of splicing polymorphisms in regulating human gene expression. *Genome Research*, *21*(4), 545–554. doi: 10.1101/gr.111211.110

Laviolette, S. R. (2007). Dopamine modulation of emotional processing in cortical and subcortical neural circuits: Evidence for a final common pathway in schizophrenia? *Schizophrenia Bulletin*, *33*(4), 971–981. doi: 10.1093/schbul/sbm048

Moltke, I., Grarup, N., Jørgensen, M. E., Bjerregaard, P., Treebak, J. T., Fumagalli, M., … Hansen, T. (2014). A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature*, *512*(7513), 190–193. doi: 10.1038/nature13425

Padilla-Coreano, N., Tye, K. M., & Zelikowsky, M. (2022). Dynamic influences on the neural encoding of social valence. *Nature Reviews Neuroscience*, *23*(9), 535–550. doi: 10.1038/s41583-022-00609-1

Pain, O., Glanville, K. P., Hagenaars, S. P., Selzam, S., Fürtjes, A. E., Gaspar, H. A., … Lewis, C. M. (2021). Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS Genetics*, *17*(5), e1009021. doi: 10.1371/journal.pgen.1009021

Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., … Lander, E. S. (2014). Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences USA*, *111*(4), E455–E464. doi: 10.1073/pnas.1322563111

Taking a lifespan approach to polygenic scores

Eloise W. Freitag[a] and Caroline M. Kelsey[b]

[a]Department of Molecular & Cellular Biology, Harvard University, Cambridge, MA, USA and [b]Department of Pediatrics, Division of Developmental Medicine, Boston Children's Hospital, Brookline, MA, USA
efreitag@college.harvard.edu
caroline.kelsey@childrens.harvard.edu

doi:10.1017/S0140525X2200245X, e215

Abstract

This commentary is a call to action for researchers to create and use genome-wide association studies (GWASs) with previously missed age groups (e.g., infancy, elderly), which will improve our ability to ask important developmental questions using genetic data to trace pathways across the lifespan.

In the target article, Burt challenges the "sociogenomics revolution," which is thriving because of the incorporation of polygenic scores (PGSs) into social science research. The ease of using PGSs is tempting; however, a source of risk exists in the confounding of genetic and environmental influences because of a variety of biological and statistical reasons. Here, we argue that there is another important confound to PGS work that was overlooked by the author: The lack of consideration of genome-wide association studies (GWASs) in the context of development. Currently, most GWASs with large sample sizes have focused on identifying risk alleles associated with outcomes in adulthood. A consequence of the lack of developmental work in this area, together with the fact that a person's genes do not change over the course of their lifetime, is the assumption that adult GWASs can and must be used to infer outcomes across the lifespan (Harden & Koellinger, 2020). This assumption becomes especially problematic when studying developmental traits, because the manifestation of underlying characteristics changes over the lifespan (Martin, Ressler, Binder, & Nemeroff, 2009).

One example of a disorder that can change in symptoms and forms across the lifespan is anxiety. Specific phobias often predominate in childhood, social anxiety increases in adolescence, panic disorder becomes more common in adulthood, and worry disorders often occur in older adults (Lenze & Wetherell, 2022). Although anxiety disorders are often comorbid and there are transdiagnostic traits shared across these anxiety disorders, there are also characteristics that are unique to each different disorder. The dynamic sets of symptoms associated with psychopathologies such as anxiety can lead to a variety of outcomes after an individual receives a diagnosis, with the associated behaviors becoming more extensive and chronic or the attenuation of symptoms leading to no longer meeting criteria for the disorder (Bystritsky, Khalsa, Cameron, & Schiffman, 2013). These diverging trajectories of

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

psychological disorders may be because of a variety of genetic and environmental factors. As a result, evidence from multiple longitudinal studies in this area supports a "developmental dynamic pattern" in which there is heterogeneity in developmental trajectories of symptoms and phenotypes across the life span (Martin et al., 2009). Using this model, as opposed to the "developmental stable model" in which genetics is thought to be mediated by one unchanging set of risk factors (Martin et al., 2009), is essential to accurately contextualize PGS studies.

The notion of dynamic genetic patterns is changing the way we approach studies of developmental traits, and this approach has been highlighted in studies on attention-deficit/hyperactivity disorder (ADHD) (Rovira et al., 2020), body mass index (BMI) (Couto Alves et al., 2019), and asthma (Pividori, Schoettler, Nicolae, Ober, & Im, 2019). Each of these studies calculated PGSs to investigate the genetic architecture underlying the trajectory of certain risk factors using data from infancy and childhood and found that the genes underlying these outcomes differed over time. More specifically, ADHD possesses a different set of genes that predict the onset and persistence of the disorder (Faraone & Larsson, 2019); BMI possesses heterogeneity at the LEPR/LEPROT gene, revealing longitudinal variation in BMI for infants versus children (Couto Alves et al., 2019); and asthma shows age-related changes across multiple points in the genome (23 genes were childhood-onset specific, one was adult-onset specific, and 37 were related to both childhood- and adult-onset asthma) (Pividori et al., 2019). These findings highlight the importance of the inclusion of wider age populations in this line of work to gain a holistic understanding of the biology underlying developmental outcomes (Couto Alves et al., 2019; Pividori et al., 2019; Rovira et al., 2020). By including age as a covariate, we can map the pathways by which genetic risk manifests across development, and thus study more effectively how various environments and interventions moderate early behavioral manifestations of risk across developmental stages (Dick et al., 2018).

Moreover, the lack of developmental work and provision of age metadata within GWASs (similar to what is being done for sex assigned at birth; Liu, Schaub, Sirota, & Butte, 2012) represents a missed opportunity to ask important questions related to stability and change in biological underpinnings of disorders over time. Other related metadata features can be similarly explored to answer major questions in the developmental field. Using developmentally informed PGSs allows us to ask important questions related to differential susceptibility, such as exploring how the interaction between low socioeconomic status and polygenic risk predicts mental health outcomes or enhancing our understanding of the effects of prenatal supplements on children's mental development (Colombo et al., 2004; Morgan, Shaw, & Olino, 2012). Deepening our understanding of previously well-established biological connections and other developmentally dynamic processes, such as epigenetics, holds promise as an illuminating direction for the field (Shulman & Elkon, 2021). In practice, this can help inform the development of interventions and treatments for individuals with genetic disorders or genetic risk factors (Dick et al., 2018). This work is a call to action for researchers to create and use GWASs with previously missed populations (e.g., early and late in life), which will improve our ability to ask important developmental questions and have a better understanding of how and why certain phenotypes change across the lifespan.

Acknowledgments. We would like to thank Drs. Charles A. Nelson and Michelle Bosquet Enlow for their helpful feedback on the commentary.

Financial support. This work was supported by the National Institute of Child Health and Human Development (F32 HD105312-01A1; to C.M.K.).

Competing interest. None.

References

Bystritsky, A., Khalsa, S. S., Cameron, M. E., & Schiffman, J. (2013). Current diagnosis and treatment of anxiety disorders. *Pharmacy and Therapeutics*, *38*(1), 30.

Colombo, J., Kannass, K. N., Jill Shaddy, D., Kundurthi, S., Maikranz, J. M., Anderson, C. J., … Carlson, S. E. (2004). Maternal DHA and the development of attention in infancy and toddlerhood. *Child Development*, *75*(4), 1254–1267.

Couto Alves, A., De Silva, N. M. G., Karhunen, V., Sovio, U., Das, S., & Taal, H. R., … Early Growth Genetics (EGG) Consortium (2019). GWAS on longitudinal growth traits reveals different genetic factors influencing infant, child, and adult BMI. *Science Advances*, *5*(9), eaaw3095.

Dick, D. M., Barr, P. B., Cho, S. B., Cooke, M. E., Kuo, S. I. C., Lewis, T. J., … Su, J. (2018). Post-GWAS in psychiatric genetics: A developmental perspective on the "other" next steps. *Genes, Brain and Behavior*, *17*(3), e12447.

Faraone, S. V., & Larsson, H. (2019). Genetics of attention deficit hyperactivity disorder. *Molecular Psychiatry*, *24*(4), 562–575.

Harden, K. P., & Koellinger, P. D. (2020). Using genetics for social science. *Nature Human Behaviour*, *4*(6), 567–576.

Lenze, E. J., & Wetherell, J. L. (2022). A lifespan view of anxiety disorders. *Dialogues in Clinical Neuroscience*, *13*, 381–399.

Liu, L. Y., Schaub, M. A., Sirota, M., & Butte, A. J. (2012). Sex differences in disease risk from reported genome-wide association study findings. *Human Genetics*, *131*(3), 353–364.

Martin, E. I., Ressler, K. J., Binder, E., & Nemeroff, C. B. (2009). The neurobiology of anxiety disorders: Brain imaging, genetics, and psychoneuroendocrinology. *Psychiatric Clinics*, *32*(3), 549–575.

Morgan, J. K., Shaw, D. S., & Olino, T. M. (2012). Differential susceptibility effects: The interaction of negative emotionality and sibling relationship quality on childhood internalizing problems and social skills. *Journal of Abnormal Child Psychology*, *40*(6), 885–899.

Pividori, M., Schoettler, N., Nicolae, D. L., Ober, C., & Im, H. K. (2019). Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: Genome-wide and transcriptome-wide studies. *The Lancet Respiratory Medicine*, *7*(6), 509–522.

Rovira, P., Demontis, D., Sánchez-Mora, C., Zayats, T., Klein, M., Mota, N. R., … Ribasés, M. (2020). Shared genetic background between children and adults with attention deficit/hyperactivity disorder. *Neuropsychopharmacology*, *45*(10), 1617–1626.

Shulman, E. D., & Elkon, R. (2021). Genetic mapping of developmental trajectories for complex traits and diseases. *Computational and Structural Biotechnology Journal*, *19*, 3458–3469.

Polygenic risk scores cannot make their mark on psychiatry without considering epigenetics

Diane C. Gooding[a,b] and Anthony P. Auger[a,c,d,e,f]

[a]Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA; [b]Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA; [c]Neuroscience Training Program, University of Wisconsin-Madison, Madison, WI, USA; [d]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, USA; [e]Department of Endocrinology and Reproductive, University of Wisconsin-Madison, Madison, WI, USA and [f]Physiology, and Department of Cellular and Molecular Pharmacology, University of Wisconsin-Madison, Madison, WI, USA

dgooding@wisc.edu; https://drdianecgooding.com

apauger@wisc.edu; https://augerlab.labs.wisc.edu

doi:10.1017/S0140525X22002412, e216

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Abstract

We generally agree with Burt's thesis. However, we note that the author did not discuss epigenetics, the study of how the environment can alter gene structure and function. Given epigenetic mechanisms, the utility of polygenic risk scores (PRS) is limited in studies of development and mental illness. Finally, in this commentary we expand upon the risks of reliance upon PRSs.

Burt's target article had many strengths, including acknowledgment of the challenges associated with reliance upon polygenic risk scores (PRSs). We concur that sociogenomics cannot and should not occur without consideration of the environment. Although Burt raises the importance of the environment, she did not mention epigenetics. Indeed, epigenetics, the study of how the environment can alter gene structure and function, is important to acknowledge. Epigenetic processes involve any change to chromatin without necessarily changing the underlying DNA code (Auger & Auger, 2011, 2013; Cuarenta et al., 2021). Epigenetic modifications typically occur through DNA methylation, histone modification, or through regulation by noncoding RNA.

One can think of DNA methylation as making a mark, whereas DNA demethylation is removing a mark (see Fig. 1A; Auger & Auger, 2013). These processes can be modified by environmental cues across development. Importantly, early life stress/adversity can have lasting consequences on DNA methylation, histone modifications, and noncoding RNA that can have a subtle or dramatic impact on an organism's health or behavior (Fig. 1B). Such a modification might occur early in life but not affect gene expression until later in life (Auger & Auger, 2017).

Another, more recent, mechanism for epigenetic modifications is via altered retrotransposon activity. Retrotransposons are autonomous elements capable of self-replication; long interspersed element 1 (LINE 1) is a retrotransposon present and presumably nonactive in humans, nonhuman primates, and rodents (Cuarenta et al., 2021). Using rats, Cuarenta et al. (2021) demonstrated that exposure to early life stress (i.e., predator odor exposure) altered LINE 1 levels and copy number within a brain region critical for juvenile social play. This suggests that early life stress can actually result in changes to DNA sequences within the brain. Converging evidence from rodent and human postmortem studies indicates that an organism's experience can not only reshape the structure and function of DNA via epigenetic modifications, but can also result in changes to DNA sequences. The regulation of DNA sequences, structure, and function by our personal or social experiences across our lifespan is generally ignored in PRSs (Fig. 1C).

PRSs are calculated by adding up the cumulative effects of each risk allele (multiplied by the effect size of each variant) to provide an index of genetic risk for a given disease (Burt, target article; Palk, Dalvie, de Vries, Martin, & Stein, 2019; Torkamani, Wineinger, & Topol, 2018). PRSs may have limited utility if they are unable to allow for epigenetic changes, in addition to the types of gene–environment interplay exemplified by gene–environment interactions (G × E).

That is, epigenetic mechanisms could influence the human neuroepigenome across time and/or at different stages across an individual's lifespan. Because of epigenetics, there is greater plasticity within our genome via alterations to the underlying DNA sequence itself in response to environmental challenges during development. There are stable epigenetic events (e.g., X inactivation and imprinting) as well as plastic epigenetic events (e.g., gene regulation). PRSs do not consider either type of event. PRSs are limited in accounting for the ways by which experience may induce variations in the genome sequence, such as gene × environment × time interactions, which are especially important in brain development (Auger & Auger, 2013).

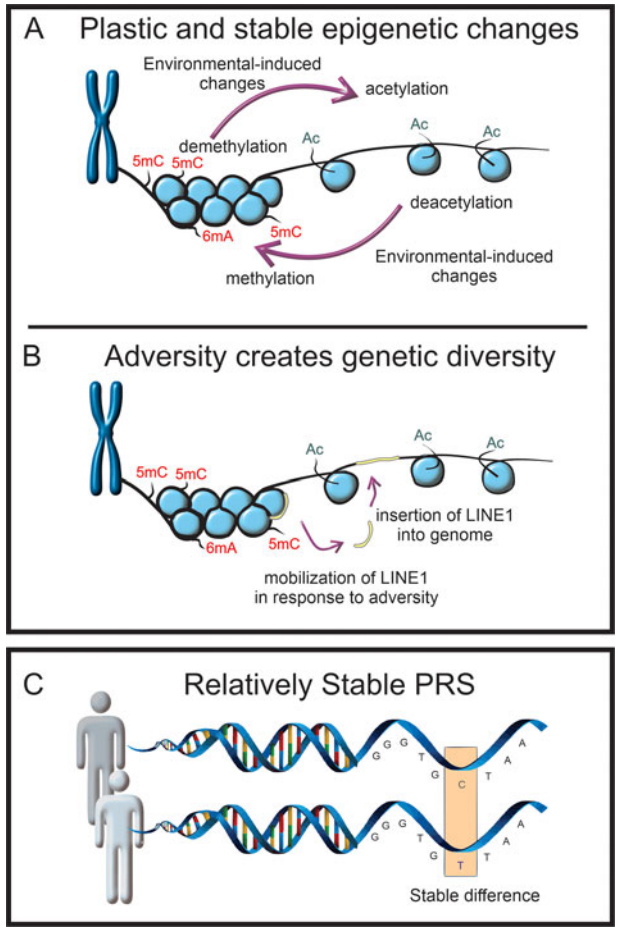

Figure 1. (Gooding and Auger). Schematic representation of how epigenetic modifications regulate gene expression. (A) Plastic and stable epigenetic changes. Environmental signals can alter DNA methylation patterns (e.g., methylation of cytosine or adenine; 5 mC and 6 mA, respectively), as well as modifications to histone (e.g., acetylation of histone tails; Ac) to impact gene expression. Generally, methylation of DNA decreases gene expression by tightening up chromatin making it inaccessible to transcription factors; whereas acetylation of histones changes the charge of histone–DNA interactions, loosening chromatin to allow DNA more assessable to transcription factors. These epigenetic modifications impacting gene expression can last hours, months, years, or be somewhat permanent. (B) Adversity creates genetic diversity. Cartoon depicting how adverse events can epigenetically modify chromatin, resulting in the activation and mobilization of the retrotransposable element, LINE 1. Once LINE 1 becomes active, it results in transcription of LINE 1 RNA, which produces critical proteins that aid its insertion back into the genome someplace else. These insertions result in increased genetic copy number of LINE 1 throughout the genome disrupting and altering gene expression in somewhat permanent ways impacting mental health and behavior. (C) Relatively stable polygenic risk scores (PRSs). This figure depicts how the individual variations in gene sequences are relatively stable, that is, not generally altered by changes in social or other environmental events even though health and behavioral consequences can be observed at individual or population levels because of modifications to the epigenome.

It is also important to consider the use of PRSs in psychiatry. When considering risks for psychiatric disorders, it is imperative that we also consider the developmental stage of the organism (Gooding & Iacono, 1995). Moreover, sex

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

differences in epigenetic mechanisms may underlie observed gender differences in prevalence, age of onset, and course of disorders such as schizophrenia and depression. PRS prediction may remain useful at the population level yet be unhelpful for assisting individuals in making predictions and prognoses. Below we use the example of schizophrenia, an exemplar of an epigenetic disorder (Gottesman, 1991; Gottesman, Shields, & Hanson, 1982).

Schizophrenia is a genetically mediated neurodevelopmental disorder characterized by etiological and phenomenological heterogeneity (Gooding, 2022; Tandon, Nasrallah, & Keshavan, 2009). Changes in LINE 1 DNA copy number have been implicated in schizophrenia (Bedrosian, Quayle, Novaresi, & Gage, 2018; Doyle et al., 2017; Jahangir, Li, Zhou, Lang, & Wang, 2022; Li et al., 2018). Studies of offspring of schizophrenia patients (e.g., Dworkin, Lewis, Cornblatt, & Erlenmeyer-Kimling, 1994; Glatt, Stone, Faraone, Seidan, & Tsuang, 2006; Gooding, Zahn-Waxler, Light, Kestenbaum, & Erlenmeyer-Kimling, 2018; Schiffman et al., 2004) suggest that impaired social functioning and emotional withdrawal in mid-childhood are predictors of schizophrenia. Recall that LINE 1 perturbations are associated with reduced social play in juvenile rodents. At present, family history (i.e., having a first-degree relative with schizophrenia) remains a more powerful predictor than a PRS (Sandstrom, Sahiti, Pavlova, & Uher, 2019).

We also recognize the potential scientific costs of reliance upon PRSs given the limited ancestral data upon which genome-wide association studies (GWASs) are based. To date, the majority of GWASs are based upon populations of European ancestry. Environmental stressors may affect different ancestral groups differentially. The disproportionate representation of European ancestry groups limits the extent to which findings can be extrapolated, as genetic prediction accuracy is substantially lower for groups of non-European ancestry. Reliance on prediction scores that are less informative in already underrepresented groups such as those of African-descent only serves to further health and healthcare disparities (Martin et al., 2019; Palk et al., 2019; Torkamani et al., 2018). Furthermore, evidence suggests that effects of environmental stressors may cause epigenetic changes that are inherited ancestrally, that is, adverse stimuli may directly affect the organism, their offspring prenatally, and future generations through epigenetic modification of the germ line (Auger & Auger, 2017; Yehuda & Lehrner, 2018; see Fig. 2). If epigenetic

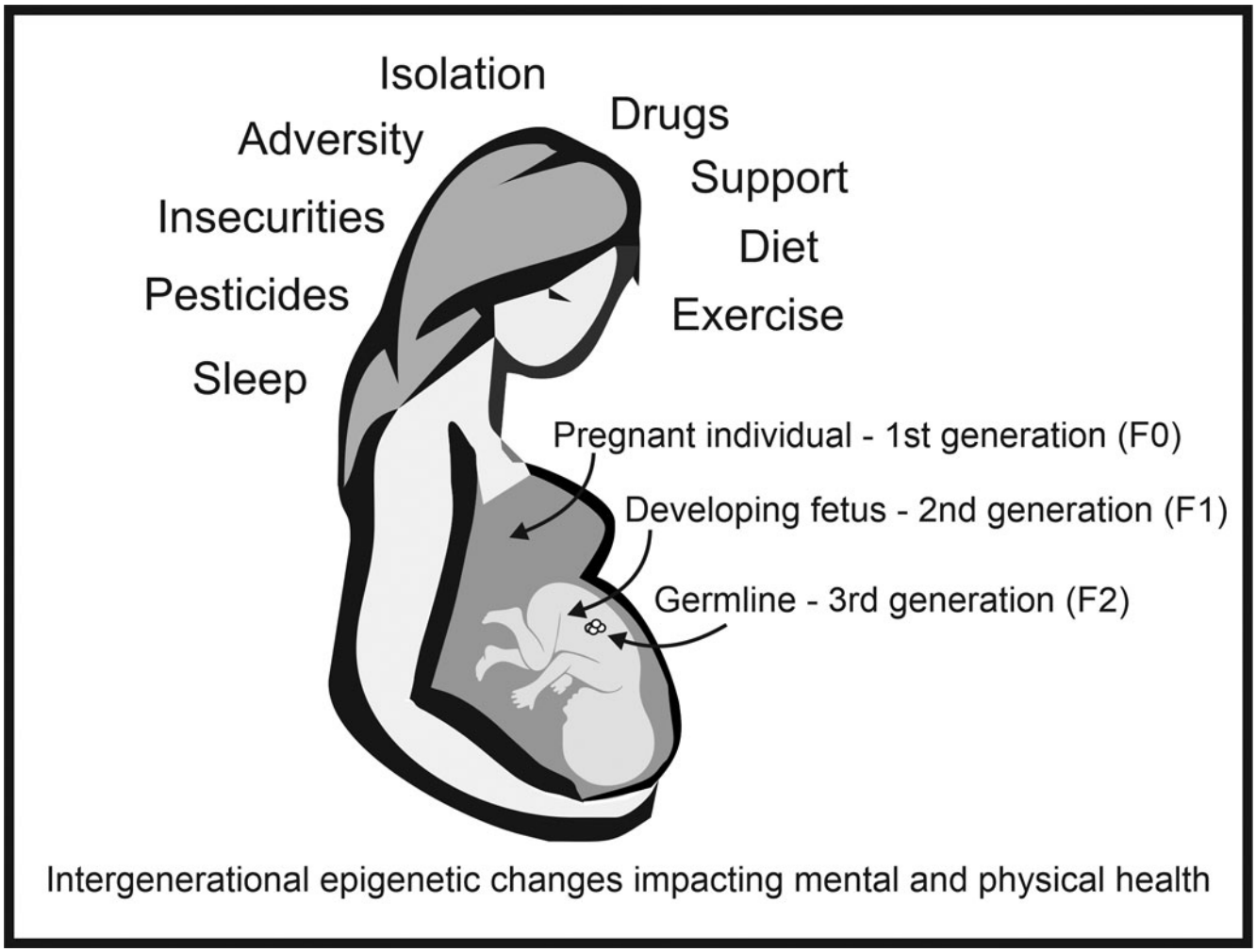

Figure 2. (Gooding and Auger). Intergenerational epigenetic changes impacting mental and physical health. A cartoon depicting how different experiences can impact the epigenome of an individual but also how these experiences can impact the epigenome of future generations. If the exposed individual (F0) is pregnant, the developing fetus (F1) is also exposed to the same events. Less considered is that in the developing F1 fetus, the germline (F2) for the subsequence generation is most likely formed and thereby is also exposed to the same perturbations. Thereby large-scale societal or individual events that impact our behavior are likely to persist in the epigenome for generations.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

mechanisms are involved in DNA perturbations that occur across generations, PRSs would be rendered less accurate.

In summary, we agree with Burt's conclusion that PRSs do not add much to our understanding of behavior in a social context. Although PRSs have some utility on a population level for predicting some health risks, we assert that reliance upon biomarkers, which can be accurately measured and reassessed following intervention, would be a more prudent guide for clinical, personal, or family decision making.

Financial support. The authors received no funding in connection with this commentary.

Competing interest. None.

References

Auger, A. P., & Auger, C. J. (2011). Epigenetic turn ons and turn offs: Chromatin reorganization and brain differentiation. *Endocrinology 152*, 349–353. https://doi.org/10.1210/en.2010-0793

Auger, A. P., & Auger, C. J. (2017). Epigenetic mechanisms shaping the brain: Implications for psychological science. In J. Call (Ed.), *APA handbook of comparative psychology, Vol. 1. Basic concepts: Methods, neural substrates, and behavior* (pp. 449–471). American Psychological Association. https://doi.org/10.1037/0000011-022

Auger, C. J., & Auger, A. P. (2013). Permanent and plastic epigenesis in neuroendocrine systems. *Frontiers in Neuroendocrinology*, *34*, 190–197. https://doi.org/10.1016/j.yfrne.2013.05.003

Bedrosian, T. A., Quayle, C., Novaresi, N., & Gage, F. H. (2018). Early life experience drives structural variation of neural genomes in mice. *Science (New York, N.Y.)*, *359*, 1395–1399.

Cuarenta, A., Kigar, S. L., Henlon, I. C., Chang, L., Bakshi, V. P., & Auger, A. P. (2021). Early life stress during the neonatal period alters social play and LINE 1 during the juvenile stage of development. *Scientific Reports*, *11*, 3549. https://doi.org/10.1038/s41598-021-82953-3

Doyle, G. A., Crist, R. C., Karatas, E. T., Hammond, M. J., Ewing, A. D., Ferraro, T. N., … Berrettini, W. H. (2017). Analysis of LINE-1 elements in DNA from postmortem brains of individuals with schizophrenia. *Neuropsychopharmacology*, *42*, 2602–2611. https://doi:10.1038/npp.2017.115

Dworkin, R. H., Lewis, J. A., Cornblatt, B. A., & Erlenmeyer-Kimling, L. (1994). Social competence deficits in adolescents at risk for schizophrenia. *The Journal of Nervous and Mental Disease*, *182*(2), 103–108.

Glatt, S. J., Stone, W. S., Faraone, S. V., Seidan, L. J., & Tsuang M. T. (2006). Psychopathology, personality traits and social development of young first-degree relatives of patients with schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *35*, 573–588.

Gooding, D. C. (2022). Brave new world: Harnessing the promise of biomarkers to help solve the epigenetic puzzle. *Schizophrenia Research*, *242*, 35–41. https://doi.org/10.1016/j.schres.2022.01.020

Gooding, D. C., & Iacono, W. G. (1995). Schizophrenia through the lens of a developmental psychopathology perspective. In Cicchetti, D. & Cohen, D. J. (Eds.), *Manual of developmental psychopathology*, Vol. II. Risk, disorder, and adaptation (pp. 535–580). Wiley.

Gooding, D. C., Zahn-Waxler, C., Light S. N., Kestenbaum, C. J., & Erlenmeyer-Kimling, L. (2018). Childhood affective indicators of risk for adulthood psychopathology: The New York high-risk project findings. *Journal of Psychiatry and Brain Science*, *3*(3), 4. https://doi.org/10.20900/jpbs.20180004

Gottesman, I. I. (1991). *Schizophrenia genesis: The origins of madness*. W.H. Freeman.

Gottesman, I. I., Shields, J., & Hanson, D. R. (1982). *Schizophrenia: The epigenetic puzzle*. Cambridge University Press.

Jahangir, M., Li, L., Zhou, J.-S., Lang, B., & Wang, X.-P. (2022). L1 retrotransposons: A potential endogeneous regulator for schizophrenia. *Frontiers in Genetics*, *13*, 878508. https://doi: 10.3389/fgene.2022.878508

Li, S., Yang, Q., Hou, Y., Jiang, T., Zong, L., Wang, Z., … Zhao, C. (2018). Hypomethylation of LINE-1 elements in schizophrenia and bipolar disorder. *Journal of Psychiatric Research*, *107*, 68–72. https://doi.org/10.1016/j.jpsychires.2018.10.009

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*, 584–591. https://doi.org/10.1038/s41588-019-0379-x

Palk, A. C., Dalvie, S., de Vries, J., Martin, A. R., & Stein, D. J. (2019). Potential use of clinical polygenic risk scores in psychiatry – Ethical implications and communicating high polygenic risk. *Philosophy, Ethics, and Humanities in Medicine*, *14*, 4. doi: doi.org/10.1186/s13010-019-0073-8

Sandstrom, A., Sahiti, O., Pavlova, B., & Uher, R. (2019). Offspring of parents with schizophrenia, bipolar disorder, and depression: A review of familial high-risk and molecular genetics studies. *Psychiatric Genetics*, *29*, 160–169. https://doi.or/10.1097/YPG.0000000000000240

Schiffman, S., Walker, E., Ekstrom, M., Schulsinger, F., Sorensen, H., & Mednick, S. (2004). Childhood videotaped social and neuromotor precursors of schizophrenia: A prospective investigation. *American Journal of Psychiatry*, *161*, 2021–2027.

Tandon, R., Nasrallah, H. A., & Keshavan, M. S. (2009). Schizophrenia, "just the facts" 4. Clinical features and conceptualization. *Schizophrenia Research*, *110*, 1–23. https://doi.10.1016/j.schres.2009.03.005

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nature Reviews*, *19*, 581–590. https://doi.org/10.1038/s41576-018-0018-x

Yehuda, R., & Lehrner, A. (2018). Intergenerational transmission of trauma effects: Putative role of epigenetic mechanisms. *World Psychiatry*, *17*, 243–257.

The value of sociogenomics in understanding genetic evolution in contemporary human populations

Ze Hong[a,b]

[a]Department of Sociology, Zhejiang University, Hangzhou, Zhejiang, China and [b]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA

ze_hong@g.harvard.edu; https://kevinhong.home.blog/

doi:10.1017/S0140525X22002424, e217

Abstract

Burt's target article oddly misses the important intellectual contribution of sociogenomics to our understanding of genetic evolution in contemporary human populations. Although social scientists' immediate research agendas are often not evolutionary in nature, I call for a better appreciation of the role of sociogenomics in answering important evolutionary questions.

"Sociogenomics" has recently become a "buzzword" in the social sciences that traditionally ignored the contribution of genetic factors in human behavior and social phenomena. As Burt noted, there have been prominent proposals to incorporate genomic methods into social scientists' toolkit and better appreciate the insights generated by genetic data (Mills & Tropf, 2020). Although I am sympathetic to Burt's call for more caution in including genetic considerations in social science research, I argue that Burt ignores one important aspect to which genetic data may greatly and uniquely contribute, and that is the study of genetic evolution in contemporary societies (I am disheartened that Burt's article does not mention "evolution" a single time).

It may be worth pointing out that the term "sociogenomics" was first proposed in evolutionary biology where researchers were deeply concerned with how the genetic basis of social behavior affects evolution (Robinson, 1999; Robinson, Grozinger, & Whitfield, 2005). Of course, sociologists' focus on contemporary social issues often makes genetic evolution, which could take considerable time, seem irrelevant, but this impression is false. In fact, natural selection is very much still operating in contemporary human populations (Milot et al., 2011; Nettle & Pollet, 2008;

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Stearns, Byars, Govindaraju, & Ewbank, 2010), and some of the most influential papers on the sociogenomics of educational attainment (e.g., Beauchamp, 2016; Kong et al., 2017) are very explicit in the use of polygenic scores to understand natural selection in contemporary human populations. Beauchamp (2016), for example, shows that natural selection has been slowly favoring lower educational attainment among individuals of European ancestry in a representative US sample. Kong et al. (2017) present similar findings in a large Icelandic sample. These findings make perfect sense given the robust negative association between fertility and educational attainment (Ní Bhrolcháin & Beaujouan, 2012; Soares, 2005): The fact that individuals with higher educational attainment tend to produce fewer children means that the genetic component will decline as long as the heritability is not exhausted (though this association is sometimes mediated by socioeconomic status; see Hugh-Jones & Abdellaoui, 2022). Empirical data in the form of heritability estimates and polygenic scores thus nicely confirm our intuition and give us a sense of the rate of natural selection (regarding the genetic component) for specific traits. As such, sociogenomics is no different in principle from nonsocial genomics which has made tremendous progress in understanding how our physical and physiological traits have responded to natural selection (Guo et al., 2018).

In general, knowing the genetic architecture of a trait (including psychological and behavioral ones) that is significantly associated with fertility is indispensable for understanding how various evolutionary forces may act on the trait, the potential genetic response, and how the phenotypic expression of the trait may change in the future. In this respect, empirical studies in sociogenomics could offer crucial guidance for theoretical and simulation models. Inspired by Beauchamp and Kong et al.'s work, I have modeled the on-going natural selection of educational attainment in contemporary societies where I show that depending on how the trait is determined by genetics and environment, we may expect rather different short-term evolutionary trajectories of both the genotype and phenotype (Hong, 2020). This type of work is necessarily provocative in the current socio-political climate, but I suggest that the genetic and cultural evolution of human behavior are both meaningful and worthy scientific endeavors, and the fact that psychological and behavioral traits are the result of complex interactions between genes and environment should not scare researchers away; rather, in the age of drastic cultural change and demographic shift (Colleran, 2016; Jensen & Levin, 2007), it is more pressing than ever to leverage insights from different disciplines to understand how fertility is associated with various traits as well as the genetic and cultural consequences, and social scientists and geneticists alike should better appreciate the value of GWASs and polygenic scores in answering important evolutionary questions.

Acknowledgments. I thank Nick Patterson for his thoughts and comments that partly inspired this commentary.

Financial support. This work was supported by the Issachar Fund and the Templeton Religion Trust.

Competing interest. None.

References

Beauchamp, J. P. (2016). Genetic evidence for natural selection in humans in the contemporary United States. *Proceedings of the National Academy of Sciences*, *113*(28), 7774–7779. https://doi.org/10.1073/pnas.1600398113

Colleran, H. (2016). The cultural evolution of fertility decline. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1692), 20150152. https://doi.org/10.1098/rstb.2015.0152

Guo, J., Wu, Y., Zhu, Z., Zheng, Z., Trzaskowski, M., Zeng, J., … Yang, J. (2018). Global genetic differentiation of complex traits shaped by natural selection in humans. *Nature Communications*, *9*(1), 1865. https://doi.org/10.1038/s41467-018-04191-y

Hong, Z. (2020). Modelling the on-going natural selection of educational attainment in contemporary societies. *Journal of Theoretical Biology*, *493*, 110210. https://doi.org/10.1016/j.jtbi.2020.110210

Hugh-Jones, D., & Abdellaoui, A. (2022). Human capital mediates natural selection in contemporary humans. *Behavior Genetics*, *52*(4-5), 1–30.

Jensen, C. J., & Levin, B. H. (2007). The world of 2020: Demographic shifts, cultural change, and social challenge. In J. Schafer (Ed.), *Exploring the future of crime, communities, and policing* (pp. 31–70). U.S. Department of Justice, Federal Bureau of Investigation.

Kong, A., Frigge, M. L., Thorleifsson, G., Stefansson, H., Young, A. I., Zink, F., … Stefansson, K. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences*, *114*(5), E727–E732. https://doi.org/10.1073/pnas.1612113114

Mills, M. C., & Tropf, F. C. (2020). Sociology, genetics, and the coming of age of sociogenomics. *Annual Review of Sociology*, *46*, 553–581. https://doi.org/10.1146/annurev-soc-121919-054756

Milot, E., Mayer, F. M., Nussey, D. H., Boisvert, M., Pelletier, F., & Reále, D. (2011). Evidence for evolution in response to natural selection in a contemporary human population. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(41), 17040–17045. https://doi.org/10.1073/pnas.1104210108

Nettle, D., & Pollet, T. V. (2008). Natural selection on male wealth in humans. *The American Naturalist*, *172*(5), 658–666.

Ní Bhrolcháin, M., & Beaujouan, É. (2012). Fertility postponement is largely due to rising educational enrolment. *Population Studies*, *66*(3), 311–327. https://doi.org/10.1080/00324728.2012.697569

Robinson, G. E. (1999). Integrative animal behaviour and sociogenomics. *Trends in Ecology and Evolution*, *14*(5), 202–205. https://doi.org/10.1016/S0169-5347(98)01536-5

Robinson, G. E., Grozinger, C. M., & Whitfield, C. W. (2005). Sociogenomics: Social life in molecular terms. *Nature Reviews Genetics*, *6*(4), 257–270. https://doi.org/10.1038/nrg1575

Soares, R. R. (2005). Mortality reductions, educational attainment, and fertility choice. *American Economic Review*, *65*(3), 580–601. https://doi.org/10.1257/0002828054201486

Stearns, S. C., Byars, S. G., Govindaraju, D. R., & Ewbank, D. (2010). Measuring selection in contemporary human populations. Nature *Reviews Genetics*, *11*(9), 611–622. https://doi.org/10.1038/nrg2831

The challenges of sociogenomics make it more, not less, worthy of careful and innovative investigation

Matthew C. Keller[a,b]

[a]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO, USA and [b]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA
matthew.c.keller@Colorado.edu

doi:10.1017/S0140525X22002291, e218

Abstract

Influences on social traits involve a tangled interplay of genetic, social, and environmental factors. Moreover, there is increasing awareness that gene–environment correlations are real and potentially measurable. Such gene–environment correlations can mislead if they are uncontrolled and genetic associations are interpreted as being purely because of direct genetic effects. This complexity is cause for more and better investigation, not a reason to refrain from researching one of the potentially important factors (genetics) influencing trait variation.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Burt casts doubt on the utility of polygenic scores (PGSs) in social science research. I applaud her attempt to avoid *ad hominem* attacks and straw man arguments. Although she largely succeeds at the former, I argue below that she sometimes fails at the latter. And although many of the limitations she describes are valid, her conclusion – that these represent fatal flaws in the use of PGSs in social science research – is not.

Burt begins and ends her paper by drawing an analogy between the excitement previously generated by the candidate gene approach and that surrounding the use of PGSs in social and behavioral genetics research. I do not believe this analogy holds. The candidate gene era laid bare the fallibility of the scientific process as currently practiced: It is likely that the many thousands of positive candidate gene findings on psychiatric and behavioral traits reported in the literature are predominated by false positives (Border et al., 2019). However, research findings on PGSs are very different. PGS findings are largely replicable and PGSs estimate true quantities. The issue at hand is how PGSs should be used and how their results should be interpreted. This is a much different and much more interesting place to be.

A core problem with Burt's critique of PGSs is that she misconstrues the state of thinking in the field, and then proceeds to argue against her own misconstruction. For example, what Burt calls "downward causation" refers to the context dependence of genetic associations, which, she argues, makes genetic associations for social behaviors "unavoidably confounded." Yet no behavioral geneticist (I hope) believes that genetic effects exist in a vacuum, independent from any potential environmental context. Alleles that influence smoking may often have different effects depending on public policy and availability of tobacco, and alleles that influence skin pigmentation probably have different influences on vitamin D sufficiency in societies that differ in average sun exposure, because of clothing or climate. This does not make these genetic effects "artificial" – it simply means they are mediated by environmental factors – nor does it make them "unavoidably confounded" – mediation and confounding are conceptually distinct. Such mediation makes PGSs more, not less, interesting, and would reduce enthusiasm for studying PGSs only if one expects that "true" genetic effects should be invariant across context. (Whether genetic effects differ across extant environmental differences is, of course, an empirical question.)

Similarly, Burt points out that simplifying assumptions made in genome-wide association studies (GWASs) (e.g., that single-nucleotide polymorphisms [SNPs] have only additive effects) or in construction of some PGSs (e.g., that all SNPs are causal) are wrong, or that the approach (e.g., only estimating effects at common variants) ignores important information. She uses these observations to imply that these approaches are therefore naïve or produce untrustworthy results. However, models are not meant mirror reality – to be so would not only be impossible but would render them incomprehensible. Models intentionally simplify to be understandable and/or to allow parameter estimation. Contrary to Burt's black-or-white thinking on this, at issue is the *degree* to which results are biased and whether this bias matters with respect to the question being investigated. It is well understood that PGSs underestimate total trait heritability, mostly because of the finite sizes of GWAS discovery samples (Wray et al., 2013). Depending on the question at hand, the underestimation may often be irrelevant (e.g., a hypothesis test of whether a depression PGS predictive ability is moderated by stressful life events; Colodro-Conde et al., 2018) or be corrected for in the model (e.g., using structural equation modeling on PGSs within families to estimate parental influences; Balbona, Kim, & Keller, 2021). The imperfect predictive ability of PGSs has no necessary relationship to their utility.

More centrally, Burt argues that PGSs of social traits are likely to be biased because of indirect genetic effects (e.g., passive gene–environment correlation or assortative mating) or because of confounding with environmental differences (e.g., as a result of uncontrolled population stratification). This arguably may not matter much if the goal of the PGS is purely to predict (Plomin & von Stumm, 2022), but it certainly matters if the goal is explanation. Again, however, this is problematic only to the extent that (a) the confounding effects exist and have not been corrected (which will inevitably occur to some degree, depending on the trait and design), and (b) that the results are interpreted as being solely because of direct genetic influences. It should be noted that the issues Burt raises regarding the interpretation of PGS results apply equally to the interpretation of GWAS effect sizes, SNP-heritability, and SNP-correlation estimates. So should these approaches be used "sparingly" when studying social or behavioral outcomes, as Burt argues? I think not. Many of the interpretational issues Burt raises are real, but they are inherent to the topics of study. Understanding the causes of individual differences in, say, educational attainment is complicated business, and must involve a tangled interplay of genetic, social, and environmental factors, all mediated through multiple different channels, but this is cause for more and better investigation, not a reason to refrain from researching one of the potentially important factors (genetics) influencing educational attainment.

Burt has identified several core issues regarding the difficulty in interpreting molecular genetic estimates, but these are neither unique to PGSs nor to social/behavioral traits. How should the field move forward in light of these issues? In agreement with Burt, there should be greater care in interpreting and describing PGS results, for example, as the relationship between a trait and "PGS estimates" rather than "genetic propensity." There is increasing awareness in the field that gene–environment correlations are real and can mislead if interpreted as being purely because of direct genetic effects – driven largely by findings from sociogenomics researchers (Abdellaoui, Dolan, Verweij, & Nivard, 2022; Berg et al., 2019; Howe et al., 2022; Kong et al., 2018; Young et al., 2018). An alternative tack is to use new designs and/or data types that allow disambiguation of environmental and genetic effects. One obvious approach is to oversample close relatives in future collections of biobank style datasets. Such within-family estimates may not provide perfect estimates of direct genetic effects, but they do control for the vast majority of potentially confounding environmental influences (Howe et al., 2022).

In summary, I have a much more optimistic view of the future of PGS research in social science than does Burt, even with its imperfections and challenges. The challenges make the topic all the more worthy of careful and innovative investigation.

Financial support. M.C.K. was supported by NIH grants MH100141 and MH130448.

Competing interest. None.

References

Abdellaoui, A., Dolan, C. V., Verweij, K. J. H., & Nivard, M. G. (2022). Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature Genetics*, *54*, 1345–1354. https://doi.org/10.1038/s41588-022-01158-0

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Balbona, J. V., Kim, Y., & Keller, M. C. (2021). Estimation of parental effects using polygenic scores. *Behavior Genetics*, *51*(3), 264–278.
Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., … Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *eLife*, *8*, e39725. https://doi.org/10.7554/eLife.39725
Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry*, *176*(5), 376–387. https://doi.org/10.1176/appi.ajp.2018.18070881
Colodro-Conde, L., Couvy-Duchesne, B., Zhu, G., Coventry, W. L., Byrne, E. M., Gordon, S., … Martin, N. G. (2018). A direct test of the diathesis–stress model for depression. *Molecular Psychiatry*, *23*(7), 1590–1596. https://doi.org/10.1038/mp.2017.130
Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., … Palviainen, T. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592.
Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science (New York, N.Y.)*, *359*(6374), 424–428. https://doi.org/10.1126/science.aan6877
Plomin, R., & von Stumm, S. (2022). Polygenic scores: Prediction versus explanation. *Molecular Psychiatry*, *27*(1), 49–52. https://doi.org/10.1038/s41380-021-01348-y
Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*, *14*(7), 507–515. https://doi.org/10.1038/nrg3457
Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., … Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*(9), 1304–1310. https://doi.org/10.1038/s41588-018-0178-9

The social stratification of population as a mechanism of downward causation

Emily Klancher Merchant

Science and Technology Studies, University of California, Davis, Davis, CA, USA
ekmerchant@ucdavis.edu
https://emilyklancher.com

doi:10.1017/S0140525X2200231X, e219

Abstract

This commentary expands on Burt's concept of downward causation to include any association between genomic variants and a given outcome that is forged through social practices rather than biochemical pathways. It proposes the social stratification of population, through which endogamy over a period of generations produces allele frequency differences between socioeconomic strata, as a mechanism of downward causation.

In "Challenging the utility of polygenic scores for social science: Environmental confounding, downward causation, and unknown biology," sociologist and criminologist Burt demonstrates that genome-wide association studies (GWASs) and polygenic scores (PGSs) for such social outcomes as educational attainment do not live up to the promises made by their enthusiasts. GWASs capture only a small portion of potential genetic influences, and PGSs are irremediably confounded by environmental factors. One often-overlooked confound is "downward causation," which Burt defines as "sociocultural forces that sort and select individuals based on genetically influenced traits, such as skin pigmentation and height, into different environments and exposures that influence social outcomes." Sociocultural sorting and selection create what Burt describes as "artificial genetic associations," or "environmental influences masquerading as genetic influences in GWAS." In this commentary, I expand the concept of downward causation to include *any association between genomic variants and a given outcome that is forged through social practices rather than biochemical pathways*. This expanded definition does not require discrimination or privilege on the basis of physiological traits; rather, it recognizes that *social stratification produces population stratification*, or what we might call *the social stratification of population*. Because of space limitations, the argument is necessarily stylized and speculative, and should be understood as suggestive rather than demonstrative.

Geneticists define population stratification as "the situation that arises when a study population contains two or more ethnic or racial subgroups that have different allele frequencies and, just coincidentally, different levels of a particular phenotype" (Hamer, 2000). Recognizing the possibility that population stratification can produce spurious associations in GWASs, geneticists usually limit GWASs to a single ancestry group, European in the case of educational attainment. However, recent studies have demonstrated that population stratification also occurs *within* ancestry groups (Haworth et al., 2019). Geneticists typically attribute this population sub-structure to micro-geographic differentials in allele frequencies. Social scientists, however, should recognize that population sub-structure may also result from social stratification: Systems of hierarchy operating within societies in which some people or groups have more status, power, and resources than others.

Although social scientists investigating epigenetics recognize that the social world can shape cellular processes (Massey et al., 2018), those using GWASs and PGSs to identify sources of individual differences in educational attainment typically assume that causality operates only in the upward direction: Genetic differences generate social hierarchy, with status, power, and resources accruing to those with genomic variants that make them more capable of success (Harden, 2021). To be sure, an individual's DNA is set at birth, and no life experience – short of high levels of toxic exposure – will change it. On the scale of historical time, however, decades of sociological research suggest the possibility of downward causation: That social hierarchy across generations can shape individual DNA. This is because, in European and Europe-descended societies, children are typically born into the same social position occupied by their parents, and a variety of institutions make mobility difficult. One such institution is endogamy – within-group marriage – which members of high-status groups use as a strategy for maintaining social hierarchy and preserving their position within it (Kalmijn, 1991; van Leeuwen, Maas, & Miles, 2005). The existence of a tightly bounded "marriage market" in the highest strata of European and American societies is so well-established and well-known as to provide the plot for an entire genre of nineteenth-century novels and such recent television shows as *Bridgerton* and *The Gilded Age*. Through several generations of endogamy, social stratification likely also contributed to the development of population sub-structure within European and Europe-descended societies.

In her article, Burt cites the now-classic "chopsticks problem" as an example of population stratification. An analogue for *the social stratification of population* would be the salad-fork problem. A GWAS conducted 100 years ago for the number of forks used per meal, even if limited to individuals of European ancestry, would likely find hits because the social groups that used the most forks per meal as a symbol of their status also had a long

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

history of using endogamy to perpetuate their status. Over generations, founder effects and genetic drift likely produced different allele frequencies between multiple-fork users and single-fork users within societies. The same variants that predicted salad-fork usage 100 years ago might also have predicted the number of servants employed or the value of property owned. Today, those same variants probably correlate with educational attainment. In the United Kingdom, it is well-known that a person's class status directly produces or precludes educational opportunities (Jackson & Marsden, 2012). Access to secondary education is more equitable in the United States, but the majority of Americans do not attend college. In the United States, the rate of high school completion grew substantially around the turn of the twentieth century. At the same time, children of the elite increasingly attended college in order to maintain and justify their status (Groeger, 2021). Because of the social stratification of population, this growing group of college graduates was likely genomically distinct from those who did not attend. As higher education expanded further over the course of the twentieth century, college graduates tended to marry one another and the college admissions process tended to privilege those whose parents graduated from college (Domingue, Fletcher, Conley, & Boardman, 2014; Neidhöfer & Stockhausen, 2019; Schwartz & Mare, 2005), further perpetuating the downward causation of single-nucleotide polymorphism (SNP)–education correlations.

Over the past 5 years, researchers in social genomics have realized that GWASs for educational attainment pick up much more than just direct genetic effects. Scientists have begun to use the term "dynastic effects" to refer to the correlation between a parent's genotype and a child's phenotype (Morris, Davies, Hemani, & Smith, 2020). This term, however, remains undertheorized and underexplored, and is often assumed to describe the direct genetic effect of the parents' genotypes on their parenting (Brumpton et al., 2020). The concept of social stratification of population, however, suggests that the PGS for educational attainment includes dynastic effects in the original sense of the term "dynasty" – a high-status lineage – reflecting the social capital marshaled by an individual's extended family.

Competing interest. None.

References

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. Å., … Davies, N. M. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nature Communications*, *11*(3519), 1–13.

Domingue, B. W., Fletcher, J., Conley, D., & Boardman, J. D. (2014). Genetic and educational assortative mating among U.S. adults. *Proceedings of the National Academy of Sciences*, *111*(22), 7996–8000.

Groeger, C. V. (2021). *The education trap: Schools and the remaking of inequality in Boston*. Harvard University Press.

Hamer, D. H. (2000). Beware the chopsticks gene. *Molecular Psychiatry*, *5*(1), 11–13.

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Haworth, S., Mitchell, R., Corbin, L., Wade, K. H., Dudding, T., Buddu-Agrey, A., … Timpson, N. J. (2019). Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications*, *10*(333), 1–9.

Jackson, B., & Marsden, D. (2012). *Education and the working class*. Routledge.

Kalmijn, M. (1991). Status homogamy in the United States. *American Journal of Sociology*, *97*(2), 496–523.

Massey, D. S., Wagner, B., Donnelly, L., McLanahan, S., Brooks-Gunn, J., Garfinkel, I., … Notterman, D. A. (2018). Neighborhood disadvantage and telomere length: Results from the Fragile Families Study. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *4*(4), 28–42.

Morris, T. T., Davies, N. M., Hemani, G., & Smith, G. D. (2020). Population phenomena inflate genetic associations of complex social traits. *Science Advances*, *6*(16), eaay0328.

Neidhöfer, G., & Stockhausen, M. (2019). Dynastic inequality compared: Multigenerational mobility in the United States, the United Kingdom, and Germany. *Review of Income and Wealth*, *65*(2), 383–414.

Schwartz, C. R., & Mare, R. D. (2005). Trends in educational assortative marriage from 1940 to 2003. *Demography*, *42*(4), 621–646.

van Leeuwen, M. H. D., Maas, I., & Miles, A. (Eds.) (2005). *Marriage choices and class boundaries: Social endogamy in history*. Cambridge University Press.

Polygenic scores ignore development and epigenetics, dramatically reducing their value

David S. Moore

Psychology Field Group, Pitzer College, Claremont, CA, USA
dmoore@pitzer.edu; http://pzacad.pitzer.edu/~dmoore/

doi:10.1017/S0140525X22002473, e220

Abstract

Polygenic scores cannot elucidate the mechanisms that produce behavioral phenotypes (including "intelligence"). Therefore, they are unlikely to yield helpful interventions. Moreover, they are poor predictors of individuals' developmental outcomes. Burt's critique is well-supported by the details of molecular biology. Specifically, experiences affect epigenetic factors that influence phenotypes via how the genome functions, a fact that lends support to Burt's conclusions.

Polygenic scores (PGSs) have excited biomedical researchers for years, but they have recently received increasing attention from social scientists interested in normal psychological variation (Harden, 2021). PGSs have been touted as a way to study "the causes and consequences" of even complex behavioral phenotypes like intelligence (Plomin & von Stumm, 2018, p. 148). However, their utility remains as controversial as that of the genome-wide association studies (GWASs) from which they derive (Charney, 2022). Burt's target article provides a potentially devastating critique of the value of these scores.

Burt correctly notes that all serious scientists accept that genetic differences "influence – in some complex, context-dependent way – developmental differences." Genetic variation is associated with phenotypic variation in part because DNA is used in the causal chain of events that builds phenotypes. Even so, this acknowledgment does not mean PGSs can offer useful predictions about individual outcomes, let alone causal insights about how to helpfully affect developmental processes.

As Turkheimer (2012) has observed, the statistical tools employed by social scientists working with genomic data have "not succeeded in discriminating actual causal processes from spurious correlations and non-causal associations" (p. 51). As a result, genomic social science can be "causally refractory… no one is about to use polygenic scores to figure out why children excel or fail in school or become addicted to drugs" (Turkheimer, 2019, p. 46). Understanding causation in ways that permit beneficial intervention requires experimental studies that are entirely unlike the correlational research that generates PGSs.

In fact, even if a researcher's goal is to predict rather than to elucidate causation, the correlations that yield many PGSs

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

cannot be trusted. For example, the enormity of the data set used by Lee et al. (2018) – which involved a sample of 1.1 *million* individuals and produced one of the most highly regarded social science PGSs to date – ensured that some *arbitrary* correlations would inevitably appear to be "significant" (see Richardson & Jones [2019] for this argument). Consequently, it is unsurprising that Morris, Davies, and Smith (2020) found educational-outcome PGSs to have predictive accuracy that is "poor…at the individual level [and] … . inferior to [that associated with] parental socioeconomic factors. [These scores] failed to accurately predict later achievement…[and] currently have limited use for accurately predicting individual educational performance" (p. 1). Likewise, Harden and Koellinger (2020) wrote "even the best currently available PGS for behavioural outcomes cannot make accurate predictions for the outcome of any specific individual" (p. 570).

Clearly, PGSs cannot be appropriately used for predicting individual outcomes. But making predictions is the best that correlational studies like GWASs can offer; because correlation does not indicate causation, GWASs cannot deliver effective treatments for behavioral challenges, either. If PGSs cannot be used to accurately predict individual outcomes and if the GWAS that gives rise to them cannot inspire effective interventions, these approaches should be understood to be of negligible value.

If the intended purpose of PGSs is to reveal something about individuals' "genetic propensities" (Harden et al., 2020, pp. 1, 2, 5), this negative assessment of their value does not merely reflect an immature state of the art. Instead, it is unlikely that PGSs will *ever* be of much value. This is because DNA segments are used differently in different contexts (Lickliter, 2017; Moore, 2001; Noble, 2006, 2012; Pan, Shai, Lee, Frey, & Blencowe, 2008; Waddington, 1957, 1968). As all phenotypes can be influenced by variable non-genetic factors, there can be no *absolute* "genetic propensity" for any phenotype, as a "propensity" in one context could very well *not* be a "propensity" in another context. Ultimately, the notion of "genetic potential" is unfounded, because genetic factors specify a *norm* of reaction, not a restricted *range* of reaction (Gottlieb, 1995); because phenotypic outcomes are not constrained by genomes that operate in context-independent ways, it will always be impossible to identify context-independent "genetic propensities." Remember, our developmental contexts are not fixed – after all, humans throughout history have continually invented new modes of education that have exposed children to never-before-experienced contexts – so a genotype that contributes to a below average phenotype in many contexts could nonetheless contribute to an above average phenotype in other not-yet-explored contexts (Lewontin, 2000). As Burt stated, "the context-specificity of PGSs… precludes their use as 'genetic potential' in general." I agree: Future refinement of PGSs will still not allow them to accurately characterize individuals' "genetic propensities."

Contexts are crucial in phenotypic development in part because they affect the epigenetic states of genomes, thereby altering how those genomes work (Moore, 2017). Burt's article draws appropriately critical conclusions regarding PGSs, but it omits mention of this important phenomenon. Although one's genetic sequence is thought to remain unchanged across the lifespan – which PGS proponents consider to be a strength of these scores – it is now clear that identical genomes can function differently depending on the experiential histories of the individual twins containing those genomes (Fraga et al., 2005; Morgan, Sutherland, Martin, & Whitelaw, 1999). As a result, sequence data alone cannot lead to accurate predictions about developmental outcomes; the mere presence, in a cell, of a DNA segment with a particular sequence will have no functional consequences if that segment has been dramatically down-regulated via epigenetic mechanisms such as DNA methylation or histone modification (Moore, 2013, 2015, 2016). And because experiential factors like social status (Tung et al., 2012), diet (Morgan et al., 1999), and maternal deprivation (Provencal et al., 2012), for example, have been experimentally shown to epigenetically change genomic activity and phenotypic outcomes in mammals, the idea that evaluating a genome at conception could provide accurate insights into much-later-developing phenotypes should be recognized as fundamentally flawed.

Twenty-first century instantiations of behavioral genetics – including GWASs and the PGSs they generate – remain targets of valid criticism (Charney, 2022; Richardson & Jones, 2019; Turkheimer, 2012). Given molecular biologists' understanding that DNA, epigenetic processes, and contextual factors work together in interdependent ways to produce phenotypes that are in no way pre-specified in the genome, these latest attempts to predict behavioral outcomes from DNA sequence information alone are bound to fail.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Charney, E. (2022). The "golden age" of behavior genetics? *Perspectives on Psychological Science*, *174*(4), 1188–1210. https://doi.org/10.1177/17456916211041602

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., … Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the USA*, *102*, 10604–10609.

Gottlieb, G. (1995). Some conceptual deficiencies in "developmental" behavior genetics. *Human Development*, *38*(3), 131–141.

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Harden, K. P., Domingue, B. W., Belsky, D. W., Boardman, J. D., Crosnoe, R., Malanchini, M., … Harris, K. M. (2020). Genetic associations with mathematics tracking and persistence in secondary school. *NPJ Science of Learning*, *5*(1), 1–8. https://doi.org/10.1038/s41539-020-0060-2

Harden, K. P., & Koellinger, P. D. (2020). Using genetics for social science. *Nature Human Behavior*, *4*, 567–576. https://doi.org/10.1038/s41562-020-0862-5

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*, 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Lewontin, R. C. (2000). *The triple helix: Gene, organism, and environment*. Harvard University Press.

Lickliter, R. (2017). Developmental evolution. *WIREs Cognitive Science*, *8*, e1422. https://doi.org/10.1002/wcs.1422

Moore, D. S. (2001). *The dependent gene: The fallacy of "Nature vs. Nurture"*. Holt.

Moore, D. S. (2013). Behavioral genetics, genetics, and epigenetics. In P. D. Zelazo (Ed.), *Oxford handbook of developmental psychology, Vol. 1: Body and mind* (pp. 91–128). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199958450.013.0005

Moore, D. S. (2015). *The developing genome: An introduction to behavioral epigenetics*. Oxford University Press.

Moore, D. S. (2016). Behavioral epigenetics. *WIREs Systems Biology and Medicine*, *9*(1), e1333. http://dx.doi.org/10.1002/wsbm.1333

Moore, D. S. (2017). The potential of epigenetics research to transform conceptions of phenotype development. *Human Development*, *60*, 69–80. http://dx.doi.org/10.1159/000477992

Morgan, H. D., Sutherland, H. G., Martin, D. I., & Whitelaw, E. (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics*, *23*, 314–318.

Morris, T. T., Davies, N. M., & Smith, G. D. (2020). Can education be personalised using pupils' genetic data? *eLife*, *9*, e49962. https://doi.org/10.7554/eLife.49962

Noble, D. (2006). *The music of life: Biology beyond genes*. Oxford University Press.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Noble, D. (2012). A theory of biological relativity: No privileged level of causation. *Interface Focus*, *2*, 55–64. https://doi.org/10.1098/rsfs.2011.0067

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high throughput sequencing. *Nature Genetics*, *40*, 14131415. https://doi.org/10.1038/ng.259

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*, 148–159. https://doi.org/10.1038/nrg.2017.104

Provencal, N., Suderman, M. J., Guillemin, C., Massart, R., Ruggiero, A., Wang, D., … Szyf, M. (2012). The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells. *Journal of Neuroscience*, *32*(44), 15626–15642. https://doi.org/10.1523/JNEUROSCI.1470-12.2012

Richardson, K., & Jones, M. C. (2019). Why genome-wide associations with cognitive ability measures are probably spurious. *New Ideas in Psychology*, *55*, 35–41. https://doi.org/10.1016/j.newideapsych.2019.04.005

Tung, J., Barreiro, L. B., Johnson, Z. P., Hansen, K. D., Michopoulos, V., Toufexis, D., … Gilad, Y. (2012). Social environment is associated with gene regulatory variation in the rhesus macaque immune system. *Proceedings of the National Academy of Sciences of the USA*, *109*, 6490–6495.

Turkheimer, E. (2012). Genome wide association studies of behavior are social science. In K. S. Plaisance & T. A. C. Reydon (Eds.), *Philosophy of behavioral biology. Boston studies in the philosophy of science* (Vol. *282*, pp. 43–64). Springer. https://doi.org/10.1007/978-94-007-1951-4_3

Turkheimer, E. (2019). The social science blues. *Hastings Center Report*, *49*(3), 45–47. https://doi.org/10.1002/hast.1008

Waddington, C. H. (1957). *The strategy of the genes*. Allen & Unwin.

Waddington, C. H. (1968). The basic ideas of biology. In C. H. Waddington (Ed.), *Towards a theoretical biology (Vol. 1: Prolegomena, pp. 1–32)* (pp. 9–10). Edinburgh University Press.

Genomics might not be the solution, but epistemic validity remains a challenge in the social sciences

David Moreau and Kristina Wiebels

School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand
d.moreau@auckland.ac.nz
k.wiebels@auckland.ac.nz
https://braindynamicslab.com/
https://kwiebels.github.io/

doi:10.1017/S0140525X22002357, e221

Abstract

We sympathize with many of the points Burt makes in challenging the value of genetics to advance our understanding of social science. Here, we discuss how recent reflections on epistemic validity in the behavioral sciences can further contribute to a reappraisal of the role of sociogenomics to explain and predict human traits, aptitudes, and achievement.

In a detailed and compelling synthesis, Burt questions the validity of genomics to gain knowledge about social science issues. Burt's cautionary note is a much-needed reminder of the limitations of sociogenomics, a field that has seen a fair amount of hype in the last few years. Here, we wish to comment on a central argument Burt makes – that because they are already well-measured behaviorally, constructs like academic achievement or cognitive aptitudes have little to benefit from the tools of sociogenomics. In our view, this argument potentially disregards the serious challenges psychologists face in measuring these constructs, no matter how well-defined they may seem behaviorally.

Psychological constructs are not pure, assumption-free operationalizations of the underlying traits or abilities they are meant to represent. Rather, the way constructs are validated and refined over time follows a process whereby convergent validity dictates the empirical instantiations – in the form of tests, tasks, or measures – that are hypothesized to probe the same constructs, and those that in contrast are understood to tap different ones. As a result, constructs are heavily influenced by initial operationalizations, in a process that is biased toward convergence at the risk of failing to explore valid – or sometimes better – alternatives (Moreau & Wiebels, 2022).

In this context, we should be cautious about uncritically ascribing validity to psychological constructs on the basis of psychometric convergence or divergence with one another. No matter how objective they might seem, behaviorally assessed constructs remain subjective and far from assumption-free. It does not follow that genomics is *necessarily* the answer to help refine our understanding of psychological constructs, but we should refrain from thinking that the measurement of constructs in the behavioral sciences is as good as it can be, or that only improvements in psychometric properties will lead to better, more valid assessments. Sociogenomics may or may not be the solution, but epistemic validity remains a challenge in the field.

Not all fields to which genomic tools are applied suffer equally from this limitation. For example, this bias is arguably less problematic when genomics is applied to medicine, where it has led to major advances in our understanding of cancer, heritable disorders, or infectious disease outbreaks (McCarthy, McLeod, & Ginsburg, 2013). Success in the clinical domain remains highly heterogeneous, however, with most significant advances having been achieved for conditions within which constructs of interest (e.g., diagnosis) are well defined and fairly objective, often because they are based on the presence or absence of biological features. In contexts where diagnosis is more subjective and constructs of interest less well defined – for example with psychiatric disorders diagnosed primarily from the presence of behaviors or related symptoms – genomic-based advances have been less prominent, for the same reasons they have been of somewhat limited benefit in psychology thus far.

So what could sociogenomics contribute to our understanding of aptitudes and achievement that current behavioral measures do not? The potential is wide-ranging and multifaceted, but one application that stands out is with respect to behavioral interventions designed to improve cognitive performance or abilities (Madole & Harden, 2023). Recent attempts to improve cognitive abilities have suffered from major setbacks, with strong initial claims failing to stand up to scrutiny (Moreau, 2022; Moreau, Macnamara, & Hambrick, 2019; Sala & Gobet, 2019). One of the main issues that has been identified is the lack of mechanistic understanding for the behavioral dynamics elicited by interventions, especially given the important heterogeneity in individual responses (Moreau, 2021). Some individuals show promising improvements post-interventions, whereas others do not appear to benefit at all, and current models provide little insight into the determinants of individual differences (Moreau, 2022). Together with efforts to improve and refine measurement in the context of interventions (Moreau & Wiebels, 2021), the field of genomics has the potential to shed light on the complex interactions at play to determine – and eventually predict – individual responses in a personalized manner. To be successful, such behavioral interventions are likely to require precision regimens,

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

whereby individual characteristics – potentially including genomic information – are leveraged to establish responder profiles and thereby determine the optimal blend for a particular person at a particular time.

Despite the potential for sociogenomics in this space, tangible progress remains dependent on addressing current limitations in the use of polygenic scores, especially issues such as confounding and stratification. Although these limitations might be alleviated in the context of interventions because of the controlled nature of these designs, they generally remain issues that the field of sociogenomics will need to grapple with. In addition, when genetic data are incorporated into intervention designs and individual response predictions, researchers should explicitly specify in what ways they can lead to qualitative improvements, and the potential downsides. Polygenic scores remain probabilistic, and as such include wide individual differences in the target trait at all levels (Plomin, DeFries, Knopik, & Neiderhiser, 2016); fair and accurate assessments of what sociogenomics can and cannot contribute at this time are to the benefit of all.

Finally, efforts to incorporate sociogenomics within behavioral interventions should not divert from attempts to address the structural scarcity and inequality inherent to systems and institutions. In particular, the notion that success is primarily driven by aptitudes or abilities that can be targeted by interventions has been shown to be problematic or even dangerous in some instances (Moreau, 2022; Nathan, 2017). When unchallenged, this view can prevent the implementation of institutional reforms that are known to effectively reduce systemic inequalities (Furnham, 2003). Addressing inequalities is an endeavor that often requires collective action on multiple fronts, and gaining a better understanding of individual differences via genomics is, albeit promising, only one of the many facets that can be leveraged to make progress in this direction.

Financial support. D.M. and K.W. are supported by a Marsden grant from the Royal Society of New Zealand and a University of Auckland Early Career Research Excellence Award.

Competing interest. None.

References

Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences*, *34*(5), 795–817.

Madole, J., & Harden, K. (2023). Building causal knowledge in behavior genetics. *Behavioral and Brain Sciences*, 1–76. https://doi.org/10.1017/S0140525X22000681

McCarthy, J. J., McLeod, H. L., & Ginsburg, G. S. (2013). Genomic medicine: A decade of successes, challenges, and opportunities. *Science Translational Medicine*, *5*(189), 189sr4.

Moreau, D. (2021). Shifting minds: A quantitative reappraisal of cognitive intervention research. *Perspectives on Psychological Science*, *16*(1), 148–160.

Moreau, D. (2022). How malleable are cognitive abilities? A critical perspective on popular brief interventions. *American Psychologist*, *77*(3), 409–423.

Moreau, D., Macnamara, B. N., & Hambrick, D. Z. (2019). Overstating the role of environmental factors in success: A cautionary note. *Current Directions in Psychological Science*, *28*(1), 28–33.

Moreau, D., & Wiebels, K. (2021). Assessing change in intervention research: The benefits of composite outcomes. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–14.

Moreau, D., & Wiebels, K. (2022). Psychological constructs as local optima. *Nature Reviews Psychology*, *1*, 188–189.

Nathan, L. (2017). *When grit isn't enough: A high school principal examines how poverty and inequality thwart the college-for-all promise*. Beacon Press.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*(1), 3–23.

Sala, G., & Gobet, F. (2019). Cognitive training does not enhance general cognition. *Trends in Cognitive Sciences*, *23*(1), 9–20.

Tractable limitations of current polygenic scores do not excuse genetically confounded social science

Damien Morris[a], Stuart J. Ritchie[a] and Alexander I. Young[b,c]

[a]Social, Genetic, and Developmental Psychiatry Centre, King's College London, London, UK; [b]UCLA Anderson School of Management, Los Angeles, CA, USA and [c]Human Genetics Department, UCLA David Geffen School of Medicine, Los Angeles, CA, USA
damien.morris@kcl.ac.uk
stuart.j.ritchie@kcl.ac.uk
AlexTISyoung@gmail.com
https://www.kcl.ac.uk/people/damien-morris
https://www.kcl.ac.uk/people/stuart-ritchie
https://geneticvariance.wordpress.com/

doi:10.1017/S0140525X22002503, e222

Abstract

Burt's critique of using polygenic scores in social science conflates the "scientific costs" of sociogenomics with "sociopolitical and ethical" concerns. Furthermore, she paradoxically enlists recent advances in controlling for environmental confounding to argue such confounding is scientifically "intractable." Disinterested social scientists should support ongoing efforts to improve this technology rather than obstructing progress and excusing genetically confounded research.

Burt promises her readers a dispassionate essay challenging the "value of polygenic scores [(PGS)]… for social science." She states she will do this "not by questioning the ethical or sociopolitical implications of this work…but by scrutinizing the science," and on this basis will conclude that the "scientific costs outweigh [the] meager benefits." But the "scientific costs" she enumerates – "obscuring environmental influences, perpetuating a flawed concept of genetic potential…and wasting resources" – are not scientific critiques at all but precisely the "sociopolitical and ethical concerns" she disavows.

Were this a disinterested critique focused on scientific accuracy, Burt would be as concerned about *exaggerating* the effects of "structural disadvantages and cultural influences" as "obscuring" them. Instead, she admits to "holding sociogenomic methods to higher standards than standard social science methodologies," excusing this double-standard on the basis of the "social risks" she pledged to leave aside. Similarly, she argues the "scientific costs" of "promoting PGS as 'genetic potential'…include…promoting the individualization of social problems." These are ideological objections, not scientific ones.

Burt cautions against "wasting finite resources searching for 'genes for educational attainment'" – that is, by performing genome-wide association studies (GWASs) that identify genetic

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

variants associated with individual differences in social science outcomes. But a substantial share of the funding for GWASs comes from private and philanthropic sources who disagree with Burt's assessment. As for the remainder, what could be more "sociopolitical" than the question of how taxpayer dollars should be directed by the government and its agencies? Besides, this puts sociogenomics in a Catch-22: Should we fund research to address some of the limitations of PGSs that Burt raises in her essay, or should we give up in despair? Burt counsels despair: "the production of environmentally confounded genetic associations with complex social outcomes is not simply a tractable empirical problem to be addressed with more sophisticated methods. Rather, such confounding is inevitable."

However, Burt's four substantive criticisms of using PGSs for behavioral outcomes – "relatedness confounding, downward causation, limited coverage of genetic influences, and context-specificity" – are scientifically tractable issues that have substantially been addressed. Within-family studies that use parent or sibling PGS as control variables largely address issues of population stratification and familial confounding, as Burt essentially acknowledges. Furthermore, constructing PGSs from within-family GWASs can remove confounding biases from the PGS. In addition, if the genetic variants associated with behavioral outcomes are principally expressed in the brain rather than in the skin, hair, or musculoskeletal system (e.g., Lee et al., 2018) this constrains the possibility that reported associations are confounded by "downward causation," that is, by "social selection on attractiveness, height, weight, [or] colorism." Conducting GWASs in large samples with whole-exome or whole-genome sequencing can increase the fraction of genetic influences covered by PGSs by capturing the effects of rarer variants and shed light on the basic biology underlying behavioral differences (Chen et al., 2022). Finally, extending GWASs to more historically, geographically, and culturally diverse samples will help to quantify the effects of different social contexts on the strength and direction of genetic associations.

The Catch-22 is, however, inescapable: "Even if the problems with environmental confounding could be solved," Burt insists, "the justification for incorporating PGS into social science is lacking." This is because, according to Burt, we know the answers to all the important questions already. We don't need PGSs "to demonstrate that supportive, stimulating parenting is associated with child educational attainment" because "we can observe and measure different …background factors and assess how these affect student progressions through educational systems." But Burt blurs the distinction between the language of association and the language of causation ("affect"). Environmental causation is precisely what genetically controlled designs help establish in observational research. And although we might not need PGSs to recognize "that children who experience childhood disadvantage are not able to fully realize their educational potential," they can help us more accurately quantify the *extent* to which various environmental disadvantages account for observed differences in social outcomes and measure how much these effects differ across contexts and conditions. Burt is keen to emphasize the context-specificity of genetic and environmental influences on social outcomes but, as one psychologist forcefully put it, using this as a pretext for "abandoning quantitative estimates is practically and theoretically bankrupt" (Rowe, 1994, p. 24).

The upshot of Burt's critique seems to be that social scientists can safely ignore genetics so long as they include a boilerplate disclaimer that "genetic differences… matter in a complex, context-sensitive way." But the extent of genetic confounding is not mysterious or unquantifiable. Although Burt is correct that using current PGSs to control for genetic influences is partial at best, a well-established literature going back to the 1970s has used genetically sensitive study designs to investigate social science outcomes. These not only include conventional twin studies (which consistently show genetic differences account for a substantial share of the observed individual differences in social science outcomes, e.g., Frisell, Pawitan, Långström, & Lichtenstein, 2012; Hyytinen, Ilmakunnas, Johansson, & Toivanen, 2019; Silventoinen et al., 2020) but also a panoply of other genetically sensitive designs, such as adoption designs, extended twin designs, sibling difference designs, and more (Baier, Eilertsen, Ystrom, Zambrana, & Lyngstad, 2022; Björklund & Salvanes, 2011; Holmlund, Lindahl, & Plug, 2011; Sariaslan et al., 2021; Wolfram & Morris, 2022). These various designs show substantially attenuated statistical associations between predictor and outcome after controlling for genetic confounds and sometimes remove the original association altogether. Burt insists her article is a broadside against "the scientific value of adding genetics to social science" generally, and not just an argument "about the value of PGS for social science," yet she neglects to explain why these older, kinship-based designs can be safely ignored.

Burt is correct that social scientists should include appropriate caveats when incorporating PGSs into their work and take efforts to control for environmental confounding. But for the reasons outlined above, they should also support ongoing scientific endeavors to improve this technology. They should not – as Burt does – use tractable limitations of research incorporating PGSs as a pretext to obstruct progress or to excuse genetic confounding in social science research.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Baier, T., Eilertsen, E. M., Ystrom, E., Zambrana, I. M., & Lyngstad, T. H. (2022). An anatomy of the intergenerational correlation of educational attainment – Learning from the educational attainments of Norwegian twins and their children. *Research in Social Stratification and Mobility*, *79*, 100691. https://doi.org/10.1016/j.rssm.2022.100691.

Björklund, A., & Salvanes, K. G. (2011). Chapter 3 – Education and family background: Mechanisms and policies. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. *3*, pp. 201–247). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00003-X

Chen, C.-Y., Tian, R., Ge, T., Lam, M., Sanchez-Andrade, G., Singh, T., … Runz, H. (2022). *The impact of rare protein coding genetic variation on adult cognitive function*. medRxiv. (p. 2022.06.24.22276728). https://doi.org/10.1101/2022.06.24.22276728

Frisell, T., Pawitan, Y., Långström, N., & Lichtenstein, P. (2012). Heritability, assortative mating and gender differences in violent crime: Results from a total population sample using twin, adoption, and sibling models. *Behavior Genetics*, *42*(1), 3–18. https://doi.org/10.1007/s10519-011-9483-0

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, *49*(3), 615–651. https://doi.org/10.1257/jel.49.3.615

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

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112. https://doi.org/10.1038/s41588-018-0147-3

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior*. Guilford Press

Sariaslan, A., Mikkonen, J., Aaltonen, M., Hiilamo, H., Martikainen, P., & Fazel, S. (2021). No causal associations between childhood family income and subsequent psychiatric disorders, substance misuse and violent crime arrests: A nationwide Finnish study of >650 000 individuals and their siblings. *International Journal of Epidemiology*, *50*(5), 1628–1638. https://doi.org/10.1093/ije/dyab099

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., … Kaprio, J. (2020). Genetic and environmental variation in educational attainment: An individual-based analysis of 28 twin cohorts. *Scientific Reports*, *10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Wolfram, T., & Morris, D. (2022). *Conventional twin studies overestimate the environmental differences between families relevant to educational attainment*. PsyArXiv. https://doi.org/10.31234/osf.io/m4eqv

Increasing the use of functional and multimodal genetic data in social science research

Benjamin C. Nephew[a], Chris Murgatroyd[b], Justin J. Polcari[a], Hudson P. Santos Jr.[c] and Angela C. Incollingo Rodriguez[d]

[a]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA, USA; [b]School of Healthcare Science, Manchester Metropolitan University, Manchester, UK; [c]Department of Nursing, University of Miami, Coral Gables, FL, USA and [d]Department of Psychological Science, Worcester Polytechnic Institute, Worcester, MA, USA

bcnephew@aol.com
c.murgatroyd@mmu.ac.uk
jjpolicari@wpi.edu
hsantos@miami.edu
acrodriguez@wpi.edu

doi:10.1017/S0140525X2200228X, e223

Abstract

Genetic studies in the social sciences could be augmented through the additional consideration of functional (transcriptome, methylome, metabolome) and/or multimodal genetic data when attempting to understand the genetics of social phenomena. Understanding the biological pathways linking genetics and the environment will allow scientists to better evaluate the functional importance of polygenic scores.

The article by Burt is timely in that it raises the importance of needing "dialogue between social and behavioral scientists about the scientific value of adding genetics to social science at the current state of knowledge." We agree with many of the issues raised and the complications of incorporating polygenic scores (PGSs) and genetics into models without fully understanding "the scientific costs." It is clear that PGSs need to be controlled for "population stratification, familial confounding, and downward (socio-environmental) causation" and the exposome of an individual but treating these solely as environmental confounders neglects the important impact of biology.

For example, though exposure to adversity at a particular timepoint increases the risk for emotional and behavioral symptoms and stress-related disorders, not everyone exposed develops these symptoms. Individuals respond to environments in different ways, and we now have methods which can provide insight into this previously "unknown biology." Although we agree that PGSs might not capture this biological risk across different studies because of many of the important points raised in this article, it does not mean that it is not valuable. PGSs and how they relate to social science need to be presented with the appropriate limitations or with additional, more functional assessments of genetic context.

Genetic studies in the social sciences could be augmented through the additional consideration of functional (transcriptome, methylome, metabolome) and/or multimodal genetic data when attempting to understand the genetics of social phenomena. PGSs should be embedded within multiple omics approaches, which could be further augmented with specific measures of physiological effectors such as functional neuroimaging or endocrine assays. Understanding the biological pathways potentially linking genetics and the environment will allow scientists to better evaluate the functional importance of PGSs.

Methylome studies have identified changes in DNA methylation as markers of overall brain health (Gadd et al., 2022). More direct epigenetic research has demonstrated links between DNA methylation and educational attainment, indicating that the methylome of lower-educated people was suggestive of exposure to pollution (van Dongen et al., 2018). DNA methylation has also been linked to chronic cannabis use with associated changes in cognitive performance (Wiedmann et al., 2022). Additional inclusion and considering of epigenetic data in PGS studies may enhance our understanding of how the environment, especially during early life, impacts our genome to induce lasting effects, allowing us to progress from environmental confounding to environmental mediation and/or modulation.

Transcriptome data have provided valuable insight into how genes play a role in complex traits and disease (Hatcher, Relton, Gaunt, & Richardson, 2019). Neuroimaging-based research has yielded associations between transcriptome-wide genes for brain structures and complex traits in different domains (Zhao et al., 2021). Cortical transcriptome changes have been specifically linked with educational attainment (Bartrés-Faz et al., 2019). Combining transcriptome data with PGSs can provide a clearer picture of which specific genes are having the most significant effects on social factors at discrete points in time. However, there are substantial challenges with transcriptome data because of tissue and temporally specific gene expression that limit its application.

The metabolome has been a topic of expanding interest in how genes affect change. Studies examining the metabolome have highlighted social-to-biological processes resulting in heath inequalities (Karimi et al., 2019). Using metabolic profiles, other investigations have revealed that social and economic factors have measurable impact on human physiology (Robinson et al., 2021). Metabolic impairment has been associated with the apolipoprotein E4 and insulin resistance in type 2 diabetes, which is often mediated by socioeconomic factors and is a major risk factor for late onset of Alzheimer's disease. These results could guide development of socioeconomic-based preventive measures and therapies for cognitive decline (Johnson et al., 2017). Similar to transcriptome data, metabolome information could provide crucial temporally specific functional insight into how environmental and social factors interact with the genome to induce change.

There is a critical need to go beyond simple PGSs and institute more comprehensive social and genetic data collection (which is becoming more readily available) to strengthen associations and

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

improve causal conclusions on how genes and environment interact to affect behavior. As has been observed with both genome- and brain-wide association studies, bigger is not always better, and an increased focus on smaller, more thoroughly characterized populations with functional genetic data will lead to stronger conclusions. A critical factor that needs to be considered in all studies is the growing awareness of the plasticity of genetic mechanisms of behavior, particularly the role of epigenetics.

The line between what was traditionally seen as genetic and environmental effects is increasingly blurred. Rather than environmental confounding, these may be epigenetic effects, and the discussion of PGSs in social science would be informed by a greater understanding of and appreciation for animal studies of behavioral genetics, where the bar for causal conclusions may be much higher. This is an especially important consideration in discussions of using PGSs, or any other type of genetic data, to control for genetic effects and focus on environmental factors. This is a problematic notion at the very least. Even if other types of genetic data are beyond the primary focus of the target article, we argue that consideration of functional genetic outputs is critical for future genetic studies in the social sciences, whether or not these data are collected in a particular PGS study.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Bartrés-Faz, D., González-Escamilla, G., Vaqué-Alcázar, L., Abellaneda-Pérez, K., Valls-Pedret, C., Ros, E., & Grothe, M. J. (2019). Characterizing the molecular architecture of cortical regions associated with high educational attainment in older individuals. *The Journal of Neuroscience*, *39*(23), 4566. https://doi.org/10.1523/JNEUROSCI.2370-18.2019

Gadd, D. A., Hillary, R. F., McCartney, D. L., Shi, L., Stolicyn, A., Robertson, N. A., … Marioni, R. E. (2022). Integrated methylome and phenome study of the circulating proteome reveals markers pertinent to brain health. *Nature Communications*, *13*(1), 4670. https://doi.org/10.1038/s41467-022-32319-8

Hatcher, C., Relton, C. L., Gaunt, T. R., & Richardson, T. G. (2019). Leveraging brain cortex-derived molecular data to elucidate epigenetic and transcriptomic drivers of complex traits and disease. *Translational Psychiatry*, *9*(1), 105. https://doi.org/10.1038/s41398-019-0437-2

Johnson, L. A., Torres, E. R. S., Impey, S., Stevens, J. F., & Raber, J. (2017). Apolipoprotein E4 and insulin resistance interact to impair cognition and alter the epigenome and metabolome. *Scientific Reports*, *7*(1), 43701. https://doi.org/10.1038/srep43701

Karimi, M., Castagné, R., Delpierre, C., Albertus, G., Berger, E., Vineis, P., … Chadeau-Hyam, M. (2019). Early-life inequalities and biological ageing: A multisystem biological health score approach in *Understanding Society*. *Journal of Epidemiology and Community Health*, *73*(8), 693. https://doi.org/10.1136/jech-2018-212010

Robinson, O., Carter, A. R., Ala-Korpela, M., Casas, J. P., Chaturvedi, N., Engmann, J., … Vineis, P. (2021). Metabolic profiles of socio-economic position: A multi-cohort analysis. *International Journal of Epidemiology*, *50*(3), 768–782. https://doi.org/10.1093/ije/dyaa188

van Dongen, J., Bonder, M. J., Dekkers, K. F., Nivard, M. G., van Iterson, M., Willemsen, G., … consortium, B. (2018). DNA methylation signatures of educational attainment. *NPJ Science of Learning*, *3*(1), 7. https://doi.org/10.1038/s41539-018-0020-2

Wiedmann, M., Kuitunen-Paul, S., Basedow, L. A., Wolff, M., DiDonato, N., Franzen, J., … Golub, Y. (2022). DNA methylation changes associated with cannabis use and verbal learning performance in adolescents: An exploratory whole genome methylation study. *Translational Psychiatry*, *12*(1), 317. https://doi.org/10.1038/s41398-022-02025-6

Zhao, B., Shan, Y., Yang, Y., Yu, Z., Li, T., Wang, X., … Zhu, H. (2021). Transcriptome-wide association analysis of brain structures yields insights into pleiotropy with complex neuropsychiatric traits. *Nature Communications*, *12*(1), 2878. https://doi.org/10.1038/s41467-021-23130-y

Cognitive traits are more appropriate for genetic analysis than social outcomes

Franck Ramus

CNRS, Ecole Normale Supérieure, PSL University, Paris, France
franck.ramus@ens.psl.eu; https://lscp.dec.ens.fr/en/member/663/franck-ramus

doi:10.1017/S0140525X22002515, e224

Abstract

The critique of the genetics of complex social outcomes is partly well-founded, insofar as social outcomes sometimes have unreliable relations with cognitive traits. But the correct conclusion is not to dismiss the entire field altogether. Rather, the implication is to redirect geneticists' attention to the stable cognitive phenotypes that are natural candidates for genetic analysis.

Burt's point that heritability estimates and polygenic scores are context- and population-dependent is well-taken and widely appreciated. However, it should not be overstated as implying that all genetic analyses are irremediably socially contingent, varying widely depending on period, culture, and context, thereby shunning any hope of identifying stable, meaningful genetic associations.

One can of course tell stories about education being something very different in a remote hunter–gather society or in the distant future, but this should not obscure the fact that the notion of educational achievement in the twenty-first century that is the current focus of genetic analysis is a well-defined and circumscribed concept that is essentially the same all over the world except for some extremely isolated cultures where schools don't exist. Even if it is true that the personality traits that were likely to attract a young woman to higher education in the 1870s and in the 2020s United States may differ to some extent, the cognitive traits (detailed further below) that would have been important for her to succeed at university in 1870 are very likely to be the same as those important in 2020, and they are also the same in the United States, in Saudi Arabia, or in Thailand, thus providing a stable basis for the genetic analysis of educational achievement. When some of these factors differ between countries or periods, this should not be cause for despair or rejection of genetic approaches, as the issue is perfectly empirically tractable: This should rather be welcomed as an opportunity to describe interesting gene–environment interactions.

Nevertheless, Burt's critique has the merit of highlighting the potential gaps between the social outcomes that are currently subjected to genetic analysis, and their cognitive basis. One should recall that social outcomes such as educational achievement or income have been genetically studied mainly because they were conveniently available in very large databases. In every genetics project, every participant answers one question about their highest obtained degree, regardless of the initial goal of the research. Thus, pooling across many projects has enabled researchers to gather the millions of participants required to compute reliable educational achievement polygenic scores (Okbay et al., 2022).

But to the cognitive scientist, this may seem a temporary distraction: These complex social outcomes are not phenotypes that

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

are under direct natural selection and that should naturally be the focus of genetic analysis. The phenotypes of interest for genetic analysis are situated at the cognitive level, where stable traits can be defined and can be the target of selection. For educational achievement, these are specific cognitive abilities: Not just general intelligence (which is itself a complex emerging property; Ramus, 2017), but its underlying components: Verbal ability, abstract reasoning, working memory, and also more specific cognitive skills such as phonological awareness (which contributes to reading acquisition) or number sense. One should not forget the popular but ill-named "noncognitive skills" (Ramus, 2022) such as conscientiousness, self-control, intrinsic motivation, grit, which do explain part of the educational achievement variance and which are also genetically influenced (Demange et al., 2021). These traits reliably underlie educational achievement regardless of time, culture, and gender of the learner, and there is every reason to think that they have a stable neural and genetic basis, which may be to a large extent similar in all populations.

Similarly, the answer to the question "have you ever had sex with someone of the same sex? Yes/No" has never been a valid phenotype for genetic analysis, but it is the one that was available for UK Biobank and 23andMe participants (Ganna et al., 2019). These authors are of course well aware that the stable cognitive trait of interest is sexual orientation, that it is continuous (e.g., as on the Kinsey scale), and that its relationship with actual sexual behaviour is imperfect, subject to social norms, to opportunities, and to many life circumstances. Genome-wide research on the genetics of sexual orientation will have to wait until an appropriate scale is rated by a sufficiently large number of participants.

An additional difficulty that may be less widely appreciated is that the cognitive functions that are under genetic influence are latent, unobservable variables, that cannot simply be equated with performance in one behavioural test. This is because any test, no matter how elementary it seems, inevitably recruits several cognitive functions. For instance, even the simplest reaction time test involves not only processing speed but also vision (or audition, to perceive the signal), sustained attention, language skills (to understand instructions), and motor skills (to produce a response). Therefore, there never is a one-to-one mapping between cognitive functions and behavioural tests. Any cognitive function can only be inferred by triangulating across several behavioural tests involving it in different ways.

This implies that research into the genetics of cognitive functions is going to be much more difficult than running a genome-wide association study (GWAS) on an answer to a single question or on a single test score. It will require administering well-designed, comprehensive test batteries to very large populations.

The conclusion is that the critique of the genetics of complex social outcomes is partly well-founded, insofar as social outcomes sometimes have unsystematic relations with cognitive traits. But the correct conclusion is not to dismiss the entire field altogether. Rather, the implication of this critique is to redirect geneticists' attention to the stable cognitive phenotypes that are natural candidates for genetic analysis. Unfortunately, studying the genetics of specific cognitive functions will take greater efforts and a longer time until the necessary test results are collected in sufficiently large genotyped populations.

Financial support. This work has received support by grants ANR-17-EURE-0017 and ANR-10-IDEX-0001-02.

Competing interest. None.

References

Demange, P. A., Malanchini, M., Mallard, T. T., Biroli, P., Cox, S. R., Grotzinger, A. D., … Nivard, M. G. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nature Genetics*, *53*(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Ganna, A., Verweij, K. J. H., Nivard, M. G., Maier, R., Wedow, R., Busch, A. S., … Zietsch, B. P. (2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science (New York, N.Y.)*, *365*(6456), eaat7693. https://doi.org/10.1126/science.aat7693

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., … Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449. https://doi.org/10.1038/s41588-022-01016-z

Ramus, F. (2017). General intelligence is an emerging property, not an evolutionary puzzle. *Behavioral and Brain Sciences*, *40*, 43–44. https://doi.org/10.1017/S0140525X1600176X

Ramus, F. (2022). *Noncognitive nonsense*. PsyArXiv. https://doi.org/10.31234/osf.io/xzv4f

Misguided model of human behavior: Comment on C. H. Burt: "Challenging the utility of polygenic scores for social science…"

Ken Richardson

Berwick upon Tweed, UK
k.richardson@mac.com

doi:10.1017/S0140525X22002333, e225

Abstract

This commentary emphasizes two problem areas mentioned by Burt. First, that within-family designs do not eradicate stratification confounds. Second, that the linear/additive model of genetic causes of form and variation is not supported by recent progress in molecular biology. It concludes with an appeal for a (biologically and psychologically) more realistic model of such causes.

Behavior geneticists tend to think that their field is unfairly controversial because of past associations with racism and eugenics. But there's more to it than that. Over the history of BG many scholars have commented on its seeming existence in a parallel universe, demanding relaxed scientific standards, building castles in the air with much reliance on "promissory notes," as Burt puts it. Regarding BG's grounding in unlikely assumptions, Kempthorne (1978, p. 18) asked "How naive can you get?" An illustration is how variations in cognition, educational attainment (EA), height, weight, and so on, are considered to be equally "complex," with similar causal patterns of form and variation, as if eons of evolution and gulfs of biological necessity had never happened.

Another example, of course, is how genome-wide association studies/polygenic scores (GWASs/PGSs) appeal to vague "phenotypes," using poorly validated measures, "surrogates" and "proxies," inferring causes from mountains of correlations that largely wash-out over time (Richardson & Norgate, 2015). Noting such thin evidential gruel Fletcher (2021, p. 256) refers to the "sleights of hand and folk wisdom from behavioral genetics." Burt expertly

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

exposes such problems in BG's contemporary search for validation at the molecular level. This commentary enlarges on two aspects.

Population stratification within-families

Burt rightly notes that the problem of population stratification haunts all interpretations of GWASs/PGSs and cannot be easily eliminated. BG researchers, however, seem to agree with Zaidi and Mathieson (cited by Burt) that "family-based studies are immune to stratification," and present the equivalent of a randomized controlled trial (RCT; Harden, 2021). However, children are not merely passive vehicles of additive G and E effects. They actively constructive perceptions of their worlds and creatively react to them.

Burt (target article, sect. 5.1.2, para. 7) mentions family dynamics that create "micro-stratification" within families. However, such dynamics also interact with wider social contexts to further generate spurious SNP–trait correlations in GWASs/PGSs. Facial appearance; height; weight; body shape; subclinical medical conditions such as myopia; hair form; and skin color (e.g., Hall, 2017), are all subject to positive/negative feedback from peers and teachers, as well as siblings and parents, acting to cultural norms. They create unequal psychological effects and reactions that course through individuals' school and occupational careers (Kraft, Kraft, Hagen, & Espeseth, 2022; Wilkinson & Pickett, 2018). Yet all will involve hundreds or thousands of SNPs covarying non-causally with psychological traits, and producing spurious GWAS/PGS results. They completely confound the preconditions for an RCT that Harden (2021) and others recommend.

Genetic causes

Burt says that "no serious scientist can suggest that genetic differences do not influence – in some complex, context-dependent way – developmental differences" (target article, sect. 6, para. 6). But we need to be clear that "influence" is not necessarily tractable as prediction. Unfortunately, prediction in BG is still dominated by the Galton/Fisher model. Despite acknowledging roles for the "environment," "interactions," and so on, mere attenuation of linear/additive genetic effects is assumed. So we get genomes described as "blueprints" (Plomin, 2019); or even as "cookbooks" (Harden, 2021); whereas Madole and Harden (2023) assert that PGSs reveal an individual's "genetic propensity for a trait" (Madole & Harden, 2023; sect. 3.1, para. 4); and that "the parental genotype causes an increase in their education" (Madole & Harden, 2023; note 4). Burt (e.g., target article, sect. 6, para. 6) reveals the fallacies, and the dangers, in such logic.

But so do waves of recent research in molecular biology. The creative, anticipatory dynamics mentioned above, at the sociocognitive level, have evolutionary precursors in learning/cognitive functions in cells, physiology, brain, and behavior (Lyon, Keijzer, Arendt, & Levin, 2021; Richardson, 2020; Shapiro, 2020). Development of form and variation does not start with gene transcription. DNA can do nothing until activated by the organism. That arises from vast signaling and metabolic networks monitoring the dynamic complexity and changeability of most natural environments. Their precursors in cytoplasm are inherited with genes, and their developmental fates are best described as emergent intelligent systems in which statistical patterns predict impending states by assimilating the covariance structures of the past and present (Richardson, 2021; Shapiro, 2020).

That fundamental, multi-level, intelligence is seen in the context-dependent recruitment of transcription factors, cofactors, enhancers, promoters, and so on (Isbel, Grand, & Schübeler, 2022). Alternative splicing produces a diversity of proteins from the same gene (Wright, Smith, & Jiggins, 2022). A single gene can be associated with the development of a variety of structures and functions (Watanabe et al., 2019). And multiple alternative pathways to desirable structural/functional endpoints are constructed in spite of genetic variation (Biddle, Martinez-Corral, Wong, & Gunawardena, 2021; Wagner & Wright, 2007).

Not even mutations comprise the random genetic lottery that Harden (2021) imagines (Monroe et al., 2022). There are also processes – what Shapiro calls natural genetic engineering (NGE) – through which intelligent cells can themselves change genetic information: "NGE is shorthand to summarize all the biochemical mechanisms cells have to cut, splice, copy, polymerize and otherwise manipulate the structure of internal DNA molecules… Totally novel sequences can result from de novo untemplated polymerization or reverse transcription of processed RNA molecules" (Shapiro, 2013, p. 287).

In other words, genes are best described as intermediary resource-providers for the organism as a whole: Servants to intelligent systems, not autonomous instructors. With the exception of relatively rare disorders or single-gene (Mendelian) variations, there are no independent "effects" of genomes. This is why Noble (2016) says we've had things the wrong way around in our descriptions of genetic causes. Baverstock (2021) uses the analogy of genes as the merchants that provide the necessary materials to build a house, but are neither the architect nor the builder. He calls for a "Copernican revolution" in our geno-centric view of living things.

Help or hindrance?

Burt asks if PGSs will ever be useful and urges caution. Fundamentally, though, it's the scientific framework – the Galton/Fisher model of heredity – that is the root of the problem. It persists, attracts funding and research effort, in spite of the logical, epistemological, and statistical errors described by Kempthorne (1978) and others, because it affirms prior socioeconomic structures. Until that model is replaced by a more biologically realistic one it will be more of a hindrance to the advancement of knowledge than a help.

Financial support. No funding agency involved.

Competing interest. None.

References

Baverstock, K. (2021). The gene: An appraisal. *Progress in Biophysics and Molecular Biology*, *164*, 46–62. https://doi.org/10.1016/j.pbiomolbio.2021.04.005

Biddle, J. W., Martinez-Corral, R., Wong, F., & Gunawardena, J. (2021). Allosteric conformational ensembles have unlimited capacity for integrating information. *eLife*, *10*, e65498. doi: 10.7554/eLife.65498

Fletcher, J. (2021). Backdoor to a dead end: A review essay. *Population and Development Review*, *48*, 253–258. https://doi.org/10.1111/padr.12487

Hall, J. C. (2017). No longer invisible: Understanding the psychosocial impact of skin color stratification in the lives of African American women. *Health & Social Work*, *42*, 71–78, https://doi.org/10.1093/hsw/hlx001

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Isbel, L., Grand, R. S., & Schübeler, D. (2022). Generating specificity in genome regulation through transcription factor sensitivity to chromatin. *Nature Reviews Genetics*, *23*(12), 728–740. https://doi.org/10.1038/s41576-022-00512-6
Kempthorne, O. (1978). Logical, epistemological and statistical aspects of nature–nurture data interpretation. *Biometrics*, *34*(1), 1–23.
Kraft, P., Kraft, B., Hagen, T., & Espeseth, T. (2022). Subjective socioeconomic status, cognitive abilities, and personal control: Associations with health behaviours. *Frontiers in Psychology*, *12*, 784758. doi: 10.3389/fpsyg.2021.784758
Lyon, P., Keijzer, F., Arendt, D., & Levin, M. (2021). Reframing cognition: Getting down to biological basics. *Philosophical Transactions of the Royal Society, B*, *376*, 20190750. https://doi.org/10.1098/rstb.2019.0750
Madole, J. W., & Harden, K. P. (2023). Building causal knowledge in behavior genetics. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X22000681
Monroe, J. G., Srikant, T., Carbonell-Bejerano, P., Becker, C., Lensink, M., Exposito-Alonso, M., … Weigel, D. (2022). Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature*, *602*, 101–105. https://doi.org/10.1038/s41586-021-04269-6
Noble, D. (2016). *Dance to the tune of life: Biological relativity*. Cambridge University Press.
Plomin, R. (2019). *Blueprint: How DNA makes us who we are*. MIT Press.
Richardson, K. (2020). In the light of the environment: Evolution through biogrammars not programmers. *Biological Theory*, *15*, 212e222.
Richardson, K. (2021). Genes and knowledge: Response to Baverstock. *Progress in Biophysics and Molecular Biology*, *167*, 12–17. doi: 10.1016/j.pbiomolbio.2021.10.003
Richardson, K. (2022). *Understanding intelligence*. Cambridge University Press.
Richardson, K., & Norgate, S. H. (2015). Does IQ really predict job performance? *Applied Developmental Science*, *19*, 153–169.
Shapiro, J. A. (2013). How life changes itself: The read-write (RW) genome. *Physics of Life Reviews*, *10*, 287–323.
Shapiro, J. A. (2020). All living cells are cognitive. *Biochemical and Biophysical Research Communications*, *564*, 134–149.
Wagner, A., & Wright, J. (2007). Alternative routes and mutational robustness in complex regulatory networks. *BioSystems*, *88*, 163–172.
Watanabe, K., Stringer, S., Frei, O., Mirkov, M. U., de Leeuw, C., Polderman, T. J. C., … Posthuma, D. (2019). A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics*, *51*, 1339–1348. doi: 10.1038/s41588-019-0481-0
Wilkinson, R., & Pickett, K. (2018). *The inner level*. Allen Lane.
Wright, C. J., Smith, C. W. J., & Jiggins, C. D. (2022). Alternative splicing as a source of phenotypic diversity. *Nature Reviews Genetics*, *23*, 697–710. https://doi.org/10.1038/s41576-022-00514-4

Methodological question-begging about the causes of complex social traits

John E. Richters

Rockville, MD, USA
john.richters@gmail.com

doi:10.1017/S0140525X22002400, e226

Abstract

Burt formulates her critique at a general level of abstraction that highlights the methodological deficiencies of sociogenomics without also calling attention to precisely the same deficiencies in the social science model she seeks to defend against its encroachments. What might have been a methodological bulwark against the excesses of sociogenomics is instead a one-sided critique that merely renews its charter.

A useful starting point for recognizing the foundational flaws of Burt's critique is the fact that her methodological case against the utility of sociogenomics research rests on a self-refuting thesis about the environmental confounding of polygenic score (PGS) associations with complex social traits. If, as Burt correctly argues, sophisticated statistical methodologies are incapable of distinguishing genetic from environmental causes of these traits, on what authority is she asserting—matter-of-factly, repeatedly, and without explanation—that "environmental effects masquerade as genetic influences" in PGS studies? How, then, can Burt know that the problem isn't instead that genetic effects masquerade as environmental influences in the empirical reports of mainstream social science? By her silence on this question-begging interpretation, Burt lays claim to the benefit of doubt where the evidence casts nothing but shadows of it. Those shadows are difficult to see at first because Burt formulates her critique at a general level of abstraction that highlights the methodological deficiencies of sociogenomics without also calling attention to exactly these same deficiencies in the social science research she seeks to defend against its encroachments. The key to recognizing these shadows and their methodological significance lies just beneath the surface of Burt's vague observation that a person's social traits "emerge from a complex interplay of environmental and genetic influences." What Burt doesn't tell us here is that the human brain is a masterpiece of complex biological design, that the functioning parameters of its psychological structures and processes are genetically underdetermined, shaped by natural selection pressures to continue evolving somatically (i.e., non-genetically) across individual lifespans in response to the adaptive demands and contingencies of everyday life (Dalton & Bergenn, 2007; Edelman & Gally, 2001; Ingold, 2008; Levin & Aharon, 2011; Mason, 2015; Richters, 1997, 2021; Waddington, 1957; Whitacre, 2010). Nor does Burt tell us that this underdetermination affords the brain astonishing *sui generis* degrees of freedom to make short- and long-term modifications to the functioning parameters of those capabilities, to acquire and create new ones, and to flexibly activate, suppress, combine, and leverage endlessly different configurations of these capacities in the service of adaptive needs. To characterize this jaw-dropping dynamic as merely a "complex interplay" is an understatement rivaling Emperor Hirohito's 1945 post-atomic bomb radio announcement to the Japanese people that "the war situation has developed not necessarily to Japan's advantage" (Frey & Eichenberger, 1991, p. 76).

Is it an accident that Burt holds back this much stronger genetic underdetermination card? That she says nothing about how underdetermination renders individuals qualitatively different from one another in terms of the functioning parameters and response dispositions of psychological structures and processes underlying their behavior? That she fails to mention that psychological heterogeneity is a pervasive, ubiquitous, defining characteristic of human functioning (Bryan, Tipton, & Yeager, 2021; McCaffrey, 2015; Moeller et al., 2022)? Or does Burt neglect to mention these things because they would call attention to something else she doesn't tell us: Namely, that the social science research she so vigorously champions but never gets around to describing is predicated on the logically implicit assumption that individuals are instead psychologically homogeneous, and that quantitative differences between them with respect to any particular pattern of overt functioning are produced by exactly the same psychological structures and processes operating in exactly the same ways in all individuals. Logically implicit is the straightforward sense that psychological homogeneity functions as the logically indispensable load-bearing support beam for a scaffolding of interdependent corollaries on which the coherence of all variable-oriented, sample-based research strategies and statistical modeling approaches to causal-theoretical inference rests, and without which they are unintelligible and

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

incoherent (Holland, 1986; Molenaar, 2004, 2015; Richters, 1997, 2021; Xie, 2011).

If the psychological homogeneity assumption and its corollaries were true, it follows that quantitative differences between individuals would reflect common underlying causes and that sophisticated statistical modeling techniques would be capable of identifying those causes in the covariance structures of aggregates. But they can't be, they don't, and they aren't. Nor, because they are predicated on the same faulty homogeneity assumption, are so-called statistical control and adjustment procedures capable of removing unwanted influences of theory-irrelevant nuisance variables from aggregate data, allowing researchers to peer through those disturbing influences for an unobstructed, as-if-by-experiment view of theory-relevant causes. Which is why the standard social science methodology Burt tells us nothing about is as intrinsically, provably, irredeemably incapable of identifying environmental causes of complex social traits as the sociogenomics methodology is incapable of distinguishing their genetic causes.

In fairness to Burt, there is no evidence in her critique that she deliberately side-steps these uncomfortable truths about social science methodology to stack the rhetorical deck in her favor. A more likely and troubling explanation is that Burt, like the vast majority of social and behavioral scientists, is genuinely unaware of the homogeneity-based foundational flaws of the standard social science model. Although easily identified through the logic of reverse engineering, the psychological homogeneity assumption is otherwise extraordinarily difficult to recognize without deliberate effort because it is so seamlessly woven into the fabric of social science research (Richters, 2021). Seamlessly enough that psychological homogeneity has flown under the radar and escaped scrutiny for the past 100 years as the root cause of psychology's notoriously slow theoretical progress, replication failures, and continuing reliance on discredited practices of null hypothesis significance testing.

Burt is right to be concerned about the overreaches of genetics enthusiasts. She also provides readers with ample justification for her concerns about the methodological deficiencies of contemporary sociogenomics. By failing to acknowledge that these legitimate concerns apply with equal force to the social science research she so vigorously defends but keeps hidden from view, however, Burt prosecutes her methodological case on an uneven playing field that belies her stated goal of establishing a foundation for meaningful dialogue about genetic and environmental influences. Equally troubling is that Burt repeatedly claims to have set aside her sociopolitical concerns about potential dangers of genetic influence claims while at the same time arguing that these dangers far outweigh the meager contribution potential of sociogenomics research and justify holding it to higher methodological standards than those of mainstream social science. What otherwise might have been a methodological bulwark against the excesses of sociogenomics is instead a question-begging, one-sided critique with far greater potential for renewing its charter.

Financial support. This commentary received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, *5*(8), 980–989.

Dalton, T. C., & Bergenn, V. W. (2007). *Early experience, the brain, and consciousness: An historical and interdisciplinary synthesis*. Psychology Press.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, *98*(24), 13763–13768.

Frey, B. S., & Eichenberger, R. (1991). Anomalies in political economy. *Public Choice*, *68*(1), 71–89.

Holland, P. W. (1986). Statistics and causal inference (in theory and methods). *Journal of the American Statistical Association*, *81*, 945–960.

Ingold, T. (2008). When biology goes underground: Genes and the spectre of race. *Genomics, Society and Policy*, *4*(1), 23.

Levin, Y., & Aharon, I. (2011). What's on your mind? A brain scan won't tell. *Review of Philosophy and Psychology*, *2*(4), 699–722.

Mason, P. H. (2015). Degeneracy: Demystifying and destigmatizing a core concept in systems biology. *Complexity*, *20*(3), 12–21.

McCaffrey, J. B. (2015). The brain's heterogeneous functional landscape. *Philosophy of Science*, *82*(5), 1010–1022.

Moeller, J., Dietrich, J., Neubauer, A. B., Brose, A., Kühnel, J., Dehne, M., & Pekrun, R. (2022). Generalizability crisis meets heterogeneity revolution: Determining under which boundary conditions findings replicate and generalize. *PsyArXiv*. https://doi.org/10.31234/osf.io/5wsna

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218.

Molenaar, P. C. (2015). On the relation between person-oriented and subject-specific approaches. *Journal of Person-Oriented Research*, *1*(1–2), 34–41.

Richters, J. E. (1997). The Hubble hypothesis and the developmentalist's dilemma. *Development and Psychopathology*, *9*(2), 193–229.

Richters, J. E. (2021). Incredible utility: The lost causes and causal debris of psychological science. *Basic and Applied Social Psychology*, *43*(6), 366–405.

Waddington, C. H. (1957). *The strategy of the genes*. Routledge.

Whitacre, J. M. (2010). Degeneracy: A link between evolvability, robustness and complexity in biological systems. *Theoretical Biology & Medical Modelling*, *7*, 6–17.

Xie, Y. (2011). Causal inference and heterogeneity bias in social science. *Information Knowledge Systems Management*, *10*(1–4), 279–289.

GWASs and polygenic scores inherit all the old problems of heritability estimates

Sahotra Sarkar

Departments of Philosophy and Integrative Biology, University of Texas at Austin, Austin, TX, USA
sarkar@austin.utexas.edu
sahotra-sarkar.org

doi:10.1017/S0140525X22002321, e227

Abstract

Polygenic score (PGS) computations assume an additive model of gene action because associations between phenotypes and alleles at different loci are compounded, ignoring interactions between alleles or loci let alone between genotype and environment. Consequently, PGSs are subject to the same objections that invalidated traditional heritability analyses in the 1970s. Thus, PGSs should not be used in the social sciences.

Burt must be strongly commended for challenging attempts to use polygenic scores (PGSs) in the social sciences. She is correct to note and emphasize the problems with any such attempt, especially those posed by haplotype–environment interactions and the unknown developmental biology of behaviorally relevant complex traits, especially in humans. However, the technical problems with

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

the construction and interpretation of PGS are much worse than what she presents.

PGSs are supposed to be indicative of the causal contribution of genes to phenotypes in individuals (Harden, 2021; Plomin & von Stumm, 2018). The first problem with PGS is that they are based on quantitative estimates of associations between alleles and phenotypes obtained from genome-wide association studies (GWASs). But these associations are notoriously population dependent: Even the same physical trait (for instance, skin, iris, or hair pigmentation in humans [Sarkar, 2021, pp. 140–142]) is associated with different sets of loci (and alleles at these loci) in different populations. Consequently, any attempt to construct a causal account from these associations must provide independent warrant for causal attributions (Woodward, 2005). None has been forthcoming even though GWASs have a multi-decade history (Sarkar, 1998, 2021).

A more important problem is that PGSs are constructed by compounding these associations between different alleles and a trait over multiple (often enough thousands of) loci. The simplest compounding strategy is to use a weighted sum but, as Burt notes, more complicated statistical compounding techniques are also routinely used. The trouble is that, although these compounding methods are designed to eliminate bias arising from non-representative sampling of genomes, none of them incorporates the biological mechanisms by which a trait is generated during organismic development from zygote to adult (in sexual organisms), that is, they ignore the mechanisms that would empirically indicate which alleles at which loci are causally most relevant. Moreover, all extant compounding methods rely on adding contributions from different alleles at each implicated locus.

Thus, the calculation of PGSs assumes an underlying linear model of gene action as did traditional heritability analysis. Because of that, they inherit all the non-additivity problems with heritability estimates that were recognized in the 1970s (Sarkar, 1998). The context then was the attempt to establish a causal connection between race and intelligence by figures such as Jensen (1969). Critics not only challenged Jensen's conclusions but also the methodology of heritability analysis on the grounds of illegitimate assumptions about the additivity of gene action that ignored interactions between alleles within loci (dominance), between loci (epistasis), and between genotype and environment. The names of these critics read like a "Who's Who" of theoretical population and quantitative genetics of the 1970s: Feldman and Lewontin (1975), Jacquard (1983), Kempthorne (1978), and Lewontin (1974).

Most importantly, Layzer (1974) analyzed in detail a causal model with the phenotype (P) being described as a mathematical function of genotype (G) and environment (E): $P = f(G, E)$ with no constraint on the functional form (f). This very general assumption is enough to show that the phenotypic variance cannot be modeled as a sum only of variances (e.g., the genotypic variance, the environmental variance, and a gene–environment interaction variance). Rather the phenotypic variance must include a large number of covariances between variables. Thus no additive model, however enhanced (as is supposedly the case for PGSs), can capture the variability of phenotypes, let alone the phenotypic values in individuals. (Additionally, in humans, the required covariances are impossible to estimate from accessible empirical data.)

Together, these results showed that: (i) heritability estimates do not allow causal inferences because of the additivity of variance problem; (ii) dependence of heritability estimates on the genotypic composition of population (which changes every generation); (iii) dependence of heritability estimates on limitations in the environments to which a population have been exposed; and (iv) dependence on interaction mechanisms such as dominance and epistasis besides those between genotype and environment. This work was synthesized in Sarkar (1998).

Post-Human Genome Project (HGP), the emergence of GWASs led to the revival of these criticisms by many prominent figures including Lander (Zuk, Hechter, Sunyaev, & Lander, 2012) and Feldman (Feldman & Ramachandran, 2018) in discussions of the so-called missing heritability problem. It was correctly pointed out that traditional heritability scores were over-estimates because of invalid additivity assumptions. Moreover, results (ii) and (ii) from the previous paragraph explain the population and context dependence of GWAS association values.

PGSs are touted as having sidestepped these problems (Plomin & von Stumm, 2018) but such claims are not credible. As noted earlier, PGS computation assumes an additive model of gene action that has been discredited by theoretical critiques of heritability analyses from the 1970s. Of course, this situation still admits the possibility that PGS values make accurate empirical predictions of phenotype but there has been no evidence produced for any such claim: Proponents of PGS use have been remarkably unwilling to make definite quantitative prediction of phenotypic values. Against this background it takes a very vivid imagination to believe that phenotypes can be determined according to an additive model of gene action that allows for no relevant interactions between alleles, loci, genotype, and environment (Sarkar, 2021). For time being PGS seems to be more akin to astrology than science: full of calculations based on no more than pious beliefs such as a commitment to genetic determinism and reductionism. The social sciences would do well to ignore PGS entirely.

Financial support. This work received no external funding.

Competing Interest. None.

References

Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. *Science (New York, N.Y.)*, *190*, 1163–1168.

Feldman, M. W., & Ramachandran, S. (2018). Missing compared to what? Revisiting heritability, genes and culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*, 20170064.

Harden, K. P. (2021). "Reports of my death were greatly exaggerated": Behavior genetics in the postgenomic era. *Annual Review of Psychology*, *72*, 37–60.

Jacquard, A. (1983). Heritability: One word, three concepts. *Biometrics*, *39*, 465–477.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1–123.

Kempthorne, O. (1978). Logical, epistemological and statistical aspects of nature–nurture data interpretation. *Biometrics*, *34*, 1–23.

Layzer, D. (1974). Heritability analyses of IQ scores: Science or numerology? *Science (New York, N.Y.)*, *183*, 1259–1266.

Lewontin, R. C. (1974). Analysis of variance and the analysis of causes. *American Journal of Human Genetics*, *26*, 400–411.

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*, 148–159.

Sarkar, S. (1998). *Genetics and reductionism*. Cambridge University Press.

Sarkar, S. (2021). *Cut-and-paste genetics: A CRISPR revolution*. Rowman & Littlefield.

Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford University Press.

Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, *109*, 1193–1198.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Beware of the phony horserace between genes and environments

Sam Trejo[a] and Daphne Oluwaseun Martschenko[b]

[a]Department of Sociology and Office of Population Research, Princeton University, Princeton, NJ, USA and [b]Stanford Center for Biomedical Ethics, Department of Pediatrics, Stanford University, Stanford, CA, USA
samtrejo@princeton.edu; www.samtrejo.com
daphnem@stanford.edu; www.daphnemartschenko.com

doi:10.1017/S0140525X22002485, e228

Abstract

Although Burt provides a valuable critique of the scientific value of integrating genetic data into social science research, she reinforces rather than disrupts the age-old horserace between genetic effects and environmental effects. We must move past this false dichotomy to create a new ontology that recognizes the ways in which genetic and environmental processes are inextricably intertwined.

Burt thoughtfully challenges the practical value of integrating molecular genetic data into social science research. In doing so, she provides a vital form of dissent that is uncommon to social and behavioral genomics – the field's critics (DOM included) more often demur the ethical and societal implications of the work. Her critique highlights the statistical and scientific limitations of current polygenic scores. She correctly emphasizes the scores' contextual and confounded nature, undermining their use as clean measures of genetic propensity (Meyer, Turley, & Benjamin, 2020; Murray, 2020) and complicating efforts to identify gene–environment interactions (Domingue, Trejo, Armstrong-Carter, & Tucker-Drob, 2020). Nonetheless, Burt makes a conceptual error when defining a genetic effect, conflating environmental mediation with environmental confounding, which ultimately leads her to an unproductive and age-old horserace between genes and the environment.

Specifically, Burt decomposes the effect of genes, as defined under the potential outcomes (or counterfactual) framework (Holland, 1986), into "upward" and "downward" sources of causation. She defines upward genetic causation as when genetic differences shape trait differences via biological pathways. For instance, she argues that the DNA related to Huntington's disease has an (upward) genetic effect on a person via their biology. In contrast, she defines downward genetic causation as when sociocultural forces act upon genetically influenced individual differences. For example, Burt argues that the areas of our genome related to skin color have only "artificial" (downward) effects on a one's life that operate through the sociocultural environment: Consider a dark-skinned girl who experiences more racial animus than her lighter complected sister and, in turn, experiences increased depressive symptoms (Laidley, Domingue, Sinsub, Harris, & Conley, 2019).

Burt is correct to point out that counterfactual thinking, the key conceptual toolkit for establishing causation (rather than mere correlation) in the social and biomedical sciences, does not distinguish between the effects of genes that do and do not operate through sociocultural pathways. However, Burt's desire to separate out genuine genetic effects from so-called artificial ones is itself built on: (1) The flawed and historically burdened idea that true effects of genes are straightforward, homogenous, and strictly biological; and (2) the misguided belief that it is possible to meaningfully distinguish between causal pathways that are, *in general*, proximal versus distal, or direct versus deeply mediated.

Burt's division of upward and downward genetic causation is well-intentioned and may be aimed at combating poor genomics communication that reinforces oversimplified and deterministic conceptions of genetic effects (Heine, 2017). For instance, consider a recent study that used a sibling design to estimate "direct individual genetic effects" on a range of traits, including depression, education, and body mass index (Howe et al., 2021). Because the effects of genes on social and behavioral traits often operate through complex causal chains that include group-level sociocultural processes, like discrimination, referring to them as "direct" or "individual" is misleading. The term "genetic effect," at present, refers to *any* causal pathway that begins with a genetic difference, allowing genes to linguistically trump any number of environmental mediators. We need new language to replace the entrenched gene–environment binary, as genetic determinist ideologies have helped establish and legitimize a wide range of social inequalities (Martschenko, Trejo, & Domingue, 2019).

However, in labeling the way genetic influences on skin color ultimately impact mental well-being as merely "artificial genetic associations," Burt makes the opposite mistake. She argues that colorism – the way in which our society discriminates based on a person's skin tone – is actually "the difference that makes a difference." Yet both colorism and genes related to skin color make a difference! In a world without colorism, inheriting different genes that influence skin tone wouldn't in turn affect a person's mental health. However, in a world with no variation in skin tone across individuals, eliminating the explicit and implicit biases that produce colorism wouldn't change depression rates.

Under the potential outcomes framework, the effects of a genetic variant and sociocultural processes are not mutually exclusive – each effect is defined by its own unique thought experiment which compares exactly two counterfactuals. One aspect of the world, the "treatment," is changed, and everything else is held the same. To say that there is a causal effect of a *genetic variant* on depression is to say that, in a world where a person was to inherit different alleles but everything else is held constant (including the way society "does" race), a person's likelihood of developing depression would change. Similarly, to say that there is a causal effect of *colorism* on depression is to say that, in a hypothetical world without the acute racism in our current world but identical in every other way (including the genetic characteristics of everyone within it), we would expect a change in the population prevalence of depression.

The effects of one's genes and one's sociocultural environment are hopelessly intertwined – indeed, each effect is defined *only* for particular states of the other. Even for Huntington's disease, how a persons' DNA ultimately affects their life is a function of environmental features, like access to long-term care and medications that help manage the symptoms. For this reason, Burt is wrong to claim that genetic sibling designs are confounded (i.e., lacking internal validity) by sociocultural

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

influences like colorism; instead, the example of skin tone and colorism highlights that such research designs identify contextual causal processes which often operate through the sociocultural features of our world (and therefore may have low external validity). We agree with her point that GWAS "cannot disentangle genetic from environmental," but the limitations are not only practical – they are conceptual. Burt's distinction between upward and downward genetic causation privileges sociocultural processes as somehow ontologically and causally prior to genetic factors, which is equally mistaken as viewing genetic factors as ontologically and causally prior to environments. Ironically, in attempting to wrest some of the counterfactual effects of genes back into the environmental fold, Burt thrusts the conversation again into a phony horserace between genes and environments, wherein opposing sides engage in a bean-counting exercise over how much outcome variation counts as genetic. We've been there before; it's an intellectual dead end.

Financial support. This work was partially supported by NHGRI grant T32HG008953 (D.O.M.) and the Center for Health and Wellbeing at Princeton University (S.T.).

Competing Interest. None.

References

Domingue, B., Trejo, S., Armstrong-Carter, E., & Tucker-Drob, E. (2020). Interactions between polygenic scores and environments: Methodological and conceptual challenges. *Sociological Science*, *7*, 365–386. https://doi.org/10.15195/v7.a19

Heine, S. J. (2017). *DNA is not destiny: The remarkable, completely misunderstood relationship between you and your genes*. Norton.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960.

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., … van der Zee, M. D. (2021). Within-sibship GWAS improve estimates of direct genetic effects. *BioRxiv*.

Laidley, T., Domingue, B., Sinsub, P., Harris, K. M., & Conley, D. (2019). New evidence of skin color bias and health outcomes using sibling difference models: A research note. *Demography*, *56*(2), 753–762. https://doi.org/10.1007/s13524-018-0756-6

Martschenko, D., Trejo, S., & Domingue, B. W. (2019). Genetics and education: Recent developments in the context of an ugly history and an uncertain future. *AERA Open*, *5*(1), 2332858418810516.

Meyer, M., Turley, P., & Benjamin, D. (2020). Genetic scoring presents opportunity, peril. The Wall Street Journal. https://medium.com/@michellenmeyer/response-to-charles-murray-on-polygenic-scores-e768cf145cc

Murray, C. (2020). Genetics will revolutionize social science. The Wall Street Journal. https://www.wsj.com/articles/genetics-will-revolutionize-social-science-11580169106

Polygenic scores and social science

Walter Veit[a] and Heather Browning[b,c]

[a]School of History and Philosophy of Science, The University of Sydney, Sydney, NSW, Australia; [b]Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK and [c]Department of Philosophy, University of Southampton, Avenue Campus, Southampton, UK
wrwveit@gmail.com; https://walterveit.com/
DrHeatherBrowning@gmail.com; https://www.heatherbrowning.net/

doi:10.1017/S0140525X22002345, e229

Abstract

It is a hotly contested issue whether polygenic scores should play a major role in the social sciences. Here, we defend a methodologically pluralist stance in which sociogenomics should abandon its hype and recognize that it suffers from all the methodological difficulties of the social sciences, yet nevertheless maintain an optimistic stance toward a more cautious use.

It is a hotly contested issue whether polygenic scores (PGSs) and genome-wide association studies (GWASs) should play a major role in the social sciences. As described in the target article, what we see is both (over)hype and a staunch opposition, with harsh accusations thrown around, straw man arguments, and *ad hominem* attacks. All this makes it difficult to not only evaluate the positions, but even to ask important methodological questions about the potential uses of these genetic tools within the social sciences.

Here, Burt offers an elegant methodological target article with the aim of addressing just this problem. In it, Burt objectively criticizes the hype that has often accompanied heritability research, without committing any of the above sins, drawing attention to the methodological limits and challenges of adding genetics research to the social sciences. Although we agree with many of Burt's points, however, we can't help but feel that she ends up overstating her conclusions and overplaying the differences between sociogenomics and traditional research within the social sciences.

In her conclusion, Burt states that "GWASs and PGSs may be powerful tools for identifying genetic associations, but they are not the right tools for understanding complex social traits" (target article, sect. 9, para. 3). Naturally, we wholeheartedly agree. However, our reasons for accepting this claim aren't a belief that these tools cannot at all help us to understand genetic influences or social outcomes, but rather that there is no such thing as *the* right tools for understanding complex social traits. That is, we do not think that there is some kind of unique or privileged combination of scientific tools for investigation of whatever complex social trait we are interested in, whether that is poverty, educational attainment, or criminal behavior. Let us elaborate.

As philosophers of science (and in particular, philosophers of the social sciences) have long recognized, complex phenomena are not to be understood through the competition of various methods with the aim of finding the ideal one, but rather through use of a broad range of tools that complement each other in various ways (Mitchell, 2009; Veit, 2019, 2021; Wimsatt, 2007; Ylikoski & Aydinonat, 2014). Although there are often conflicts within scientific disciplines regarding what sets of methods, models, experiments, and the like should be employed, these often appear to be driven by "indoctrination" into the methodology of a lab and ideological disputes over the correct methods. As the saying goes: If all you learn is how to swing a hammer, all problems will start to look like nails. But from a higher-level perspective, it is precisely because of the pluralism of different methods that science has flourished. And this conclusion, we think, likewise applies to the use of GWASs and PGSs.

These methods should not act as a replacement for standard social science tools, nor should they be seen as competitors to randomized controlled trials (RCTs) that investigate environmental

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

factors. Instead, we argue that they can provide us with a useful complement for research into the main targets of the social sciences, that is: complex causal systems with great heterogeneity and no strong generalizations. Just as the study of genome-wide associations bears the danger of falsely attributing causality to observed correlations, so too does standard social science. Burt is right in her criticism of the hype around PGSs: That they are often seen as deterministic, fail to control for a wide range of potential confounds, risk reviving the unfortunate gene-culture war, and so forth. But it is possible to arrive at such a critical stance by highlighting that sociogenomics will of course suffer from all the methodological difficulties of the social sciences – causal indeterminacy, the complexity of the social world, looping effects, and so forth. Within such an alternative picture, however, sociogenomics could still play a valuable role, within its own limited sphere.

Rather than simplifying the complexity of social phenomena, we argue that sociogenomics can help us to highlight how complex and causally interdependent social phenomena truly are. That is, we can buy into the main criticisms of the usefulness of PGSs in the social sciences, without being led to the strong conclusion that sociogenomics is methodologically doomed. Rather than returning to old and unhelpful discussions of social versus genetic causes, we think that sociogenomics might in fact help us toward a recognition of the complexity of our social traits and their myriad bases. This is how one should understand the argument that PGSs may improve RCTs by finding further variables to be controlled for (Harden, 2021). It's an embrace of a supplementary and pluralistic stance in the face of complexity. Rather than eliminating sociogenomics, or buying into the mistaken hype that it is going to replace and revolutionize standard social science, we can see its role instead as a complementary method to be added to the vast toolkit of the social sciences. Burt rightly points out that the methods as they are currently used too often fail to appreciate their own limitations, but this can be used as a starting-point, with these careful criticisms forming the basis for refining and strengthening the methods to better fit the contexts of use.

We therefore think that neither the majority of advocates nor the majority of critics of PGSs hold an adequate epistemic stance toward their use in the social sciences. Instead, we have here advocated for something of a mid-level approach, in which proponents of sociogenomics are urged to recognize the methodological difficulties of social science research and familiarize themselves with the philosophy of the social sciences in order to improve their own methods. Once the hype dies down, what remains will be better science, one practiced with adequate attention paid to the current problems and limitations of the methods. At the moment, without knowing exactly how this will unfold, we would like to avoid making any firm predictions regarding the likely payoffs of sociogenomics; however, we hold a (cautiously) optimistic stance regarding its future use.

Financial support. W.V.'s research was supported under Australian Research Council's Discovery Projects funding scheme (project number FL170100160).

Competing Interest. None.

References

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Mitchell, S. D. (2009). *Unsimple truths: Science, complexity, and policy*. University of Chicago Press.

Veit, W. (2019). Model pluralism. *Philosophy of the Social Sciences*, *50*(2), 91–114.

Veit, W. (2021). Model diversity and the embarrassment of riches. *Journal of Economic Methodology*, *28*(3), 291–303.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Harvard University Press.

Ylikoski, P., & Aydinonat, N. E. (2014). Understanding with theoretical models. *Journal of Economic Methodology*, *21*(1), 19–36.

Vertical pleiotropy explains the heritability of social science traits

Charley Xia[a,b,*] and W. David Hill[a,b,*]

[a]Lothian Birth Cohort studies, University of Edinburgh, Edinburgh, UK and [b]Department of Psychology, University of Edinburgh, Edinburgh, UK
Charley.Xia@ed.ac.uk; https://www.ed.ac.uk/profile/dr–charley–xia
David.Hill@ed.ac.uk; https://www.ed.ac.uk/profile/david–hill

doi:10.1017/S0140525X22002382, e230

Abstract

We contend that social science variables are the product of multiple partly heritable traits. Genetic associations with socioeconomic status (SES) may differ across populations, but this is a consequence of the intermediary traits associated with SES differences also varying. Furthermore, genetic data allow social scientists to make causal statements regarding the aetiology and consequences of SES.

Burt describes the signal captured by a polygenic score (PGS) derived from a genome-wide association study (GWAS) on social science traits such as education as being "artificial" and a product of social differences rather than genetic processes. As an example of downward causation, Burt provides the thought experiment posed by Jencks et al. (1972) where, in a hypothetical scenario, red-headed individuals are denied access to an education.

We argue that, just as a PGS captures the aggregate effect of each individual single-nucleotide polymorphism (SNP) used in its construction, each SNP from a GWAS conducted on education captures the aggregate effect of each heritable trait associated with differences in education. This process, referred to as vertical pleiotropy (also known a mediator variable) describes incidences where phenotype A (e.g., intelligence) is associated with phenotype B (education) and so a genetic variant found to be associated with phenotype A will also be associated with phenotype B (Fig. 1).

In Burt's hypothetical example, red hair would emerge as an intermediary phenotype between genetic inheritance and phenotypic consequence but in real data, childhood intelligence ($r_g = 0.72$, SE = 0.09) (Hill, Davies, Liewald, McIntosh, & Deary, 2016), health ($r_g = 0.56$, SE = 0.03) (Hill et al., 2019b), attention-deficit/hyperactivity disorder (ADHD) ($r_g = -0.54$, SE = 0.03) (Hill et al., 2019b), and neuroticism ($r_g = -0.23$, SE = 0.02) (Hill et al., 2020) show consistent and substantial genetic correlations with education and give an indication as to what

*Both authors contributed equally.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

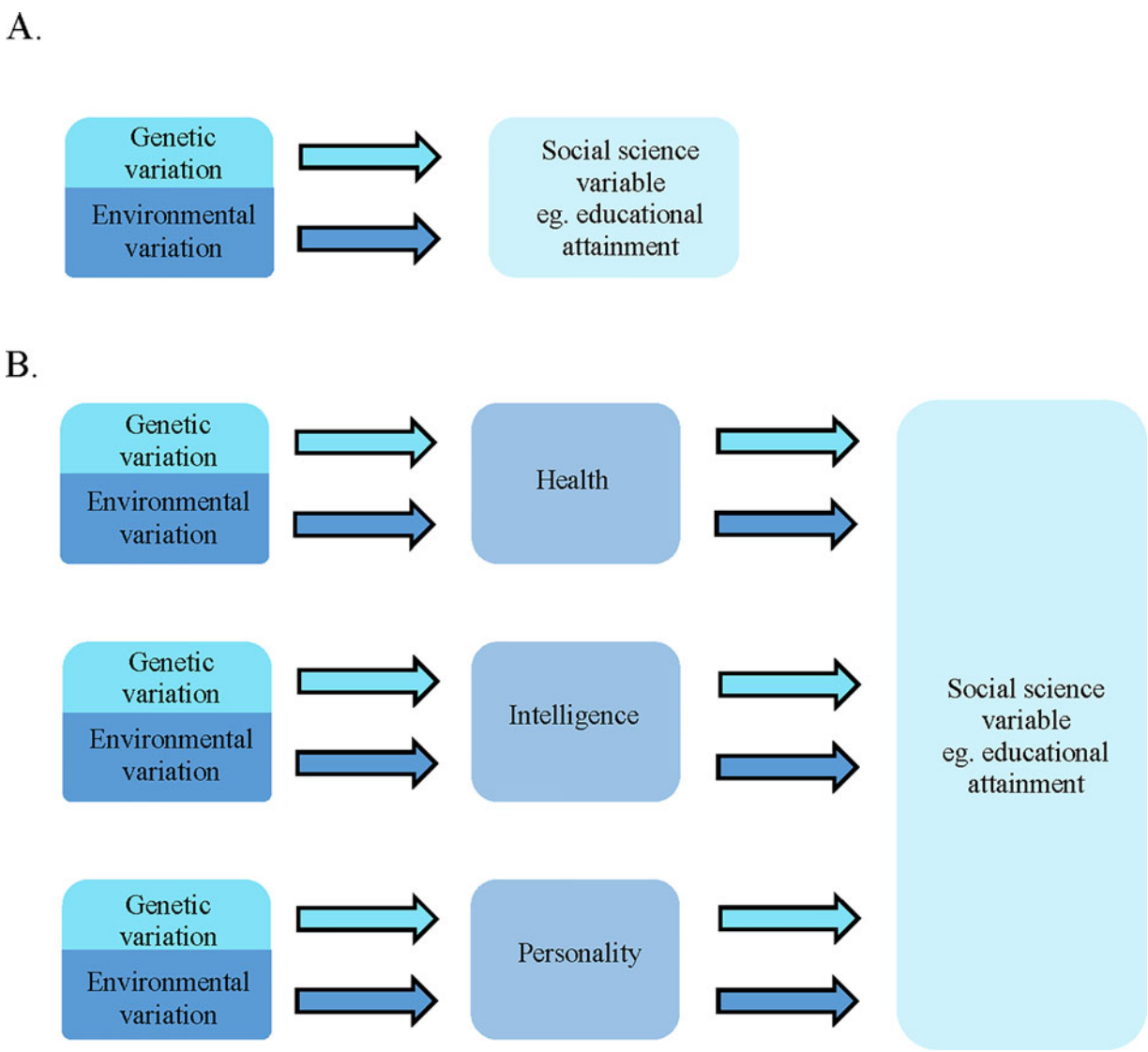

Figure 1 (Xia and Hill). Simplified illustration of vertical pleiotropy showing a subset of the possible intermediary phenotypes between genetic variation and phenotypic differences in social science variables. Illustrated is a schematic describing that when a genome-wide association study (GWAS) is performed on, or a polygenic score (PGS) is derived to predict differences in, education, genetic variation is linked to education (panel A). However, the means by which an association occurs is that, in part, a number of partly heritable traits are themselves associated with education as part of a phenotype pathway starting with genetic inheritance and ending with phenotypic consequences for education (panel B). Light blue boxes indicate sources of genetic variation whereas light blue arrows show the association between genetic and trait variation measured using GWAS or PGS. Dark blue boxes show sources of environmental variation with dark blue arrows indicating environmental associations with a trait. Pale blue boxes indicate education as an example of a social science variable. The blue/grey boxes in panel B show possible intermediary heritable phenotypes.

heritable traits may contribute towards educational attainment. In a multivariate analysis examining the traits that contribute towards education in children, Krapohl et al. (2014) found that intelligence, self-efficacy, school environment, home environment, personality, wellbeing, behavioural problems, and health, collectively explained 75% of the heritability of education.

Vertical pleiotropy also illustrates why some PGSs are population specific. When applied to education, a PGS would be population specific insofar as the heritable traits underlying educational attainment are not universal. An example of this was provided by Rimfeld et al. (2018) who showed that a PGS predicted 6.1% of education in post-Soviet Estonia compared with 2.1% in Soviet era Estonia. Furthermore, the total heritability of education in post-Soviet Estonia was estimated to be 37% compared to the Soviet era estimate of 17%. Height was used as a control variable and no significant differences between the heritability estimates were found. These differences were attributed to the rise of a more meritocratic society following the fall of the Soviet Union where individual differences in hard work and ability, which are partly genetically mediated, became the traits predictive of educational success rather than environmentally driven privilege or discrimination.

Some of the heritable traits underlying differences in education may indeed be population specific, as indicated by population-specific genetic effects on education (Rimfeld et al., 2018; Tropf et al., 2017). However, meta-analyses of GWASs of education do facilitate loci discovery, which is indicative that some of the association signal is replicated across samples and is consistent with the idea that similar heritable traits underlie

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

education differences across, predominantly European, countries and cultures.

Finally, Burt asks what the added value is of including genetics in a social science study. Mendelian randomisation (MR) is a technique that, at its heart, uses vertical pleiotropy to examine if two traits (such as, e.g., health and education) are causally connected. This is achieved by using genetic variants (such as single or multiple SNPs from a GWAS) as instrumental variables for risk factors that affect the health of a population. As genetic variants are fixed at conception their use as instrumental variables can overcome some types of confounding.

Applied to social science variables, MR has helped to understand the causes and consequences of socioeconomic status (SES) differences where intelligence has been putatively shown to be a causal factor for income (Hill et al., 2019a) and education (Anderson et al., 2020; Davies et al., 2019), where bi-directional casual effects exist in the case of education. When applied in a multivariable analysis, MR has indicated that education, and not the highly correlated trait of intelligence, is a causal factor in smoking (Sanderson, Davey Smith, Bowden, & Munafò, 2019). Conversely, higher intelligence, and not education, has been indicated to be a causal protective factor against Alzheimer's disease (Anderson et al., 2020). Using a within-family design an increase in BMI was identified as causally associated with lower levels of education (Howe et al., 2022).

In conclusion, PGS and GWAS conducted on social science traits capture the partly heritable traits that likely contribute to some of the variance of SES. Such associations are as authentic as those that act in biological pathways influencing disease traits, the difference being that, for social science traits, SNP associations are at the start of a phenotypic pathway beginning at molecular genetic inheritance and ending at phenotypic consequence. This pathway can differ between populations, but it is a strength of the molecular genetic design that MR can be applied to examine which heritable traits are causally linked to SES differences across and between cultures.

Financial support. C.X. and W.D.H. are supported by a Career Development Award from the Medical Research Council (MRC) (MR/T030852/1) for the project titled "From genetic sequence to phenotypic consequence: genetic and environmental links between cognitive ability, socioeconomic position, and health." This work was undertaken within the Lothian Birth Cohort studies group, which is supported by Age UK, the Medical Research Council (MRC), and the University of Edinburgh. For the purpose of open access, the author has applied a "Creative Commons Attribution (CC BY) licence" to any author accepted manuscript version arising from this submission.

Competing Interest. None.

References

Anderson, E. L., Howe, L. D., Wade, K. H., Ben-Shlomo, Y., Hill, W. D., Deary, I. J., … Hemani, G. (2020). Education, intelligence and Alzheimer's disease: Evidence from a multivariable two-sample Mendelian randomization study. *International Journal of Epidemiology*, *49*(4), 1163–1172.

Davies, N. M., Hill, W. D., Anderson, E. L., Sanderson, E., Deary, I. J., & Davey Smith, G. (2019). Multivariable two-sample Mendelian randomization estimates of the effects of intelligence and education on health. *eLife*, *8*, e43990. doi: 10.7554/eLife.43990

Hill, W. D., Davies, G., Liewald, D. C., McIntosh, A. M., & Deary, I. J. (2016). Age-dependent pleiotropy between general cognitive function and major psychiatric disorders. *Biological Psychiatry*, *80*(4), 266–273. https://doi.org/10.1016/j.biopsych.2015.08.033

Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., … Deary, I. J. (2019a). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*, 5741. doi: 10.1101/573691

Hill, W. D., Marioni, R. E., Maghzian, O., Ritchie, S. J., Hagenaars, S. P., McIntosh, A. M., … Deary, I. J. (2019b). A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Molecular Psychiatry*, *24*(2), 169–181. doi: 10.1038/s41380-017-0001-5

Hill, W. D., Weiss, A., Liewald, D. C., Davies, G., Porteous, D. J., Hayward, C., … Deary, I. J. (2020). Genetic contributions to two special factors of neuroticism are associated with affluence, higher intelligence, better health, and longer life. *Molecular Psychiatry*, *25*(11), 3034–3052. doi: 10.1038/s41380-019-0387-3

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., … Within Family, C. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592. doi: 10.1038/s41588-022-01062-7

Jencks, C., Smith, M., Henry, A., Jo Bane, M., Cohen, D., Gintis, H., … Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Basic Books.

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B., … Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences*, *111*(42), 15273–15278. doi: 10.1073/pnas.1408777111

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

Sanderson, E., Davey Smith, G., Bowden, J., & Munafò, M. R. (2019). Mendelian randomisation analysis of the effect of educational attainment and cognitive ability on smoking behaviour. *Nature Communications*, *10*, 2949. doi: 10.1038/s41467-019-10679-y

Tropf, F. C., Lee, S. H., Verweij, R. M., Stulp, G., van der Most, P. J., de Vlaming, R., … Mills, M. C. (2017). Hidden heritability due to heterogeneity across seven populations. *Nature Human Behaviour*, *1*(10), 757–765. doi: 10.1038/s41562-017-0195-1

Burt uses a fallacious motte-and-bailey argument to dispute the value of genetics for social science

Brendan P. Zietsch[a], Abdel Abdellaoui[b] and Karin J. H. Verweij[b]

[a]Centre for Psychology and Evolution, School of Psychology, University of Queensland, Brisbane, QLD, Australia and [b]Department of Psychiatry, Amsterdam UMC - location University of Amsterdam, Amsterdam, The Netherlands
zietsch@psy.uq.edu.au
https://psychology.uq.edu.au/profile/2404/brendan-zietsch
a.abdelllaoui@amsterdamumc.nl
https://researchinformation.amsterdamumc.org/en/persons/abdel-abdellaoui
karin.verweij@amsterdamumc.nl
https://researchinformation.amsterdamumc.org/en/persons/karin-j-h-verweij

doi:10.1017/S0140525X22002394, e231

Abstract

Burt's argument relies on a motte-and-bailey fallacy. Burt aims to argue against the value of genetics for social science; instead she argues against certain interpretations of a specific kind of genetics tool, polygenic scores (PGSs). The limitations, previously identified by behavioural geneticists including ourselves, do not negate the value of PGSs, let alone genetics in general, for social science.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

A motte-and-bailey is a fallacious argument in which an easy to defend position (motte) is conflated with a difficult to defend position (a bailey) – the latter is claimed but only the former is actually defended. Burt's article comprises such an argument. The bailey is that genetics in general is not valuable to social science. From the abstract: "Here, I challenge arguments about the value of genetics for social science," and later: "My explicit aim is to challenge the claim that genomics has much to offer social science, so much so that social science sans genetics is fatally flawed, scientifically indefensible, and possibly even morally suspect." But her actual arguments are about the limitations of one particular genetics tool, the polygenic score (PGS), and against "the claimed necessity of incorporating PGSs into social science models as measures of genetic influences." So Burt's motte is that PGSs are imperfect and should be used cautiously in social science – a position of which we know no opponents. By demonstrating this truism, she implies she has defended a much less tenable position against the value to social science of genetics in general.

First, let's establish why critics argue, and we agree, "that social science research that neglects genetics is, at best, partial and potentially flawed and misleading" (target article, sect. 1, para. 4). The fact that human behavioural traits are ubiquitously heritable (Polderman et al., 2015) – which Burt does not dispute – creates an enormous problem for social science research that ignores that fact. It means that a substantial source of individual differences remains unobserved, potentially leading to biased estimations and wrong conclusions. Any associations among different behaviours, or associations between the behaviour of parents and their children, or associations between children's behaviour and any variable influenced by parental behaviour, are likely confounded by genetic effects. Ignoring this confounding, which much social science does, renders inferences about causes of these associations invalid. For example, we might interpret the observation that children growing up with a home library have more intellectual skills as adults as a causal effect of the presence of books (Sikora, Evans, & Kelley, 2019). Or we might interpret an association between the warmth of the parent–offspring relationship during adolescence and the quality of the offspring's romantic attachments 60 years later as evidence of "the far-reaching influence of childhood environment on well-being in adulthood" (Waldinger & Schulz, 2016). The unacknowledged genetic confounds do not rule out the hypothesised causal effects, but they invalidate the evidence proffered *for* these effects (Sherlock & Zietsch, 2018).

It can therefore be vital to account for genetic confounds. PGSs are one avenue for integrating genetics into social science, but we agree with many of Burt's concerns about the usefulness and misinterpretation of PGSs, several points of which derive from our own work. We are especially concerned about the use of PGSs for "getting genetics out of the way" (target article, sect. 4.1) – that is, including a PGS in an attempt to control for genetic confounding. Isungset et al. (2022) did this in claiming to demonstrate a causal effect of parents' education on their children's school performance. They concluded that "parental educational advantage is attenuated only to a small degree when accounting for genetics." But they accounted for a PGS for educational attainment, which captures only a minority of the total genetic variance in school test scores – therefore, it is inevitable that this will only attenuate the parent–child correlation a small amount. It is invalid to infer, as the authors do, that the remaining parent–child correlation is because of a social–environmental effect of parents' education.

But this inappropriate use and interpretation of PGSs does not support Burt's argument against the value of genetics for social science. Ignoring genetics would only worsen the issue. There are various possibilities for integrating genetics into social science so as to identify, minimise, or account for genetic confounds, for example, by testing hypotheses using twin/pedigree datasets, large genetically informed (biobank) datasets, or summary-level genome-wide genetic data. Another possibility would be to adjust for the weakness of the PGS – for example, in the aforementioned Isungset et al. (2022) study the educational attainment PGS accounted for 6.3% of the variance in school test scores, whereas twin studies estimate that genetic variance accounts for ~55% of variance (Bartels, Rietveld, Van Baal, & Boomsma, 2012). Given that even accounting for this weak PGS already reduces the parent–child correlation by 14–18%, this could be consistent with complete genetic confounding of the parent–child correlation.

It might seem that there is a symmetry in Burt's arguments and ours: Burt is concerned about environmental confounding of genetic effects, whereas we are concerned about genetic confounding of environmental effects. But this leaves out important asymmetries that make Burt's overall argument unreasonable and untenable. First, while Burt argues *against* the value of genetics for social science, we argue it is important to account for both genetics *and* environmental effects, and to disentangle them where possible. Second, Burt acknowledges the great efforts that are made in genetics research to minimise the kind of environmental confounding she warns of; but on Burt's side of the debate, without taking into account genetics social science cannot minimise or even recognise genetic confounding. Third, the fixed nature of genes and well-understood process of inheritance provide natural experiments (e.g., identical and non-identical twins, Mendelian randomisation of alleles) that form the bedrock of genetics research and enable detection of genetic (and environmental) variance in traits using different analytic methodologies with different assumptions, as well as allowing cautious causal inferences. In contrast, observational/correlational research in non-genetic social science has no such avenues for establishing causality, leaving associations hopelessly confounded and making it difficult to make inferences about environmental effects.

In conclusion, Burt's argument against the value of genetics for social science is fallacious and counterproductive. The goal of understanding humans and society is best served by making the most of all available methods; accordingly, efforts should be made to integrate genetics into empirical approaches. Articles like Burt's only impede such integration.

Financial support. K.J.H.V. and A.A. are supported by the Foundation Volksbond Rotterdam.

Competing Interest. None.

References

Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2012). Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Research*, *5*(6), 544–553. doi: 10.1375/twin.5.6.544

Isungset, M. A., Conley, D., Zachrisson, H. D., Ystrom, E., Havdahl, A., Njølstad, P. R., & Lyngstad, T. H. (2022). Social and genetic associations with educational performance in a Scandinavian welfare state. *Proceedings of the National Academy of Sciences*, *119* (25), e2201869119.

Polderman, T. J., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. doi: 10.1038/ng.3285.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Sherlock, J. M., & Zietsch, B. P. (2018). Longitudinal relationships between parents' and children's behavior need not implicate the influence of parental behavior and may reflect genetics: Comment on Waldinger and Schulz (2016). *Psychological Science*, *29*(1), 154–157. doi: 10.1177/0956797617717041

Sikora, J., Evans, M. D. R., & Kelley, J. (2019). Scholarly culture: How books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Social Science Research*, *77*, 1–15. https://doi.org/10.1016/j.ssresearch.2018.10.003

Waldinger, R. J., & Schulz, M. S. (2016). The long reach of nurturing family environments: Links with midlife emotion-regulatory styles and late-life security in intimate relationships. *Psychological Science*, *27*(11), 1443–1450. doi: 10.1177/0956797616661556

Author's Response

Polygenic scores for social science: Clarification, consensus, and controversy

Callie H. Burt

Department of Criminal Justice & Criminology, Center for Research on Interpersonal Violence (CRIV), Georgia State University, Atlanta, GA, USA
cburt@gsu.edu; www.callieburt.org

doi:10.1017/S0140525X23000845, e232

Abstract

In this response, I focus on clarifying my arguments, highlighting consensus, and addressing competing views about the utility of polygenic scores (PGSs) for social science. I also discuss an assortment of expansions to my arguments and suggest alternative approaches. I conclude by reiterating the need for caution and appropriate scientific skepticism.

In my target article, I scrutinized polygenic scores (PGSs) for social science applications. Arguing that the increased uptake of PGSs in social science requires greater awareness of what PGSs are, what they measure, and how this affects their interpretation and utility, I provided an overview of PGSs with a focus on their complexities and limitations. My goal was to raise awareness of PGSs' challenges and uncertainties and promote a dialogue to foster better (social) science. I am thus grateful to the diverse group of distinguished scholars who have engaged with my article as per my aims. In 24 commentaries, scholars enriched my discussions, expanded my critiques, and/or contested my conclusions, and in so doing, raised important issues and points for fruitful debate.

The coverage in my target article coheres into two broad themes. The first concerns the challenges with PGSs I outlined, namely environmental confounding, low-resolution, and context-specificity. The second theme is the limited utility of PGSs given these challenges. I argued that the claims made by some proponents about the significant utility of PGSs for social science are overstated, even misleading. I made several recommendations, perhaps most notably that PGSs be used in social science "sparingly and cautiously with caveats placed front and center."

Commentaries largely concurred with my arguments about the limitations of PGSs. No commentator disputed my point that PGSs are not appropriately interpreted as "genetic influences" on complex social traits, as they often are. Commentators also largely agreed with my concern that PGSs are being misinterpreted or misused in some – but by no means all – sociogenomics research.[1] For example, **Keller** writes: "there should be greater care in interpreting and describing PGS results, e.g., as the relationship between a trait and 'PGS estimates' rather than 'genetic propensity'." Similarly, **Zietsch, Abdellaoui, & Verweij (Zietsch et al.)** note: "we agree with many of Burt's concerns about the usefulness and misinterpretation of PGSs, several points of which derive from our own work" (which I cited in the target article). **Fletcher** writes that "the ambiguous nature of a PGS's interpretation has led far too many investigators to over-interpret and narrowly label a PGS as 'genetic,' often to elevate the perceived importance of 'genetics' in contributing to social science outcomes." Overall, there was general consensus that researchers should not depict PGSs as reflecting "genetic influences," implicitly or explicitly.

Similarly, my explicating that the low-resolution tag-single-nucleotide polymorphism (SNP) approach of genome-wide association studies (GWASs) and PGSs, which makes them feasible, impedes their utility for gleaning biological insights was largely undisputed (but see **Alexander, Illius, Feyerabend, Wacker, & Liszkowski [Alexander et al.]**). Furthermore, no commentary challenged my argument that the context-specificity of genetic associations precludes the use of PGSs as "genetic potential" in general, and comparisons across context and condition as a means of assessing the magnitude of "genetic influences," in particular.

In contrast, my arguments about the utility of PGSs given their limitations provoked considerable debate. Some of this apparent disagreement is based on misunderstandings of either my intended arguments and recommendations or my assumptions and motivations. Importantly, genuine disagreement also exists around the tractability of the limitations and the utility of PGSs. Some commentators contend that the problems are worse that I outline and render PGSs useless, even having "negative utility" for social science (**Curtis**). Conversely, several commentators claim that the limitations with PGSs I outline are tractable and the challenges with PGSs are not as severe as I suggest.

This unique forum provides authors with the rare, valuable opportunity to immediately clarify arguments that were misunderstood and directly respond to objections. I thus devote the bulk of my comment to that end. This response is organized as follows. In section R.1, I focus on clarifying misinterpretations of my intended arguments. In section R.2, I address genuine disagreements about facts and/or their implications. Section R.3 is devoted to an assortment of commentaries that express agreement with key claims in my target article and expand my arguments in various ways. I conclude by highlighting the value of caution and appropriate scientific skepticism.

R1. Ostensible disagreements and clarifications

Several commentaries critiqued claims that I did not intend to make but that were inferred from my target article. Several of these critiques resemble or echo disputes that tend to reoccur in debates about genetics in social science and lead to tangential or misleading discussions. Thus, addressing these misunderstandings, which tend to persist, is valuable. Here, I aim to correct misconceptions that led to ostensible disagreements that do not actually exist. For clarification, I was not:

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

1.1 Opposing the use of genetics in social science, in general, or sociogenomics as a field.
1.2 Endorsing a model of psychological homogeneity or genetic sameness.
1.3 Arguing that PGSs are completely useless for social science.
1.4 Contending that all purported genetic effects on complex social traits are "artificial."
1.5 Expressing a "desire" to separate genetic from environmental influences on complex traits.
1.6 Claiming that we know all the answers to the important social science questions already.
1.7 "Vigorously defending" or "championing" social science research or measures.
1.8 Contending that a chief limitation with PGSs is that their *interpretation* is context dependent.
1.9 Challenging the use of PGSs as "genetic influences" on ideological grounds.

Assuming that these misunderstandings arose from a lack of clarity in my arguments, I address these points below. Readers who do not need this clarification may opt to skip to section R.2.

R1.1. A critique of PGSs not genetics for social science, in general

Several commentators perceived my article to be a critique of sociogenomics as a field or the incorporation of genetics into social science in any form (**Burke**; **Keller**; **Moreau & Wiebels**; **Richters**; **Zietsch et al.**). Rather than sociogenomics in general, my target article was "focus[ed] on the utility of PGSs for social science and the key premises underlying their use as measures of 'genetic propensities' for behavioral differences," as the title also announced. To be sure, I should have better worded a few sentences to reflect my specific focus on PGSs for social science; thus, I take responsibility for inadvertently encouraging this interpretation. Even so, my coverage throughout, including my key recommendations, concentrated on PGSs. This is why – to address **Morris, Ritchie, & Young's (Morris et al.)** critique – other methods of incorporating genetics into social science were not discussed. This is also why – to address **Zietsch et al.**'s primary critique – my article does not reflect "a fallacious motte-and-bailey argument" (see Shackel, 2005). My focus, which **Fletcher** aptly described as being "on a subset of 'genetics' [for social science] – the use of polygenic scores" was not a stand-in or "motte" for general opposition to genetics in social science.

Some commentaries interpret my article as implying it is acceptable to "deliberately ignor[e] genetic influences" on social phenomenon (**Burke**, also **Zietsch et al.**). Although I do not concede that genetics is relevant to the explanation of all social phenomenon (e.g., the association between being American and driving on right and being British and driving on left), my critique of PGSs was not a call for social scientists to "deliberately ignore genetics" but to recognize that however relevant genetics are to our development and social traits, PGSs do not capture "genetic (vs. environmental) influences" on social traits. By analogy, my air quality indicator is unable to accurately differentiate between carbon dioxide and volatile organic compounds (VOCs). I do not recommend you use it to measure VOCs for that reason, but from that it does not follow that I think VOCs are not important to measure, much less that they be deliberately ignored.

Similarly, **Keller** depicts me as holding a "black-and-white" position that we "should refrain from researching one of the important factors (genetics) influencing trait variation." He further implies that my arguments rest on the naïve position that models need to be perfect to be useful. Neither are positions I hold or espoused in the target article. The challenges of PGSs for social science are not merely that they are imperfect as all methods are, but rather that PGSs have specific limitations that vitiate their utility for social science research. As **Curtis** writes, articulately precising my arguments: "PGSs are so poor at capturing the genetic variation which is biologically relevant while at the same time being profoundly influenced by exactly the kind of confounders social scientists do not want contaminating their research such as race, socioeconomic status and parental characteristics."

R1.2. Not assuming psychological homogeneity

In a response familiar to critics of behavior genetics, **Richters** alleges that I, likely ignorantly, endorse a model of "psychological homogeneity" (see, Harden [2021] for an analogous "genetic sameness" argument). I do not (e.g., Burt, 2020; Simons & Burt, 2011). My scrutinizing methodological tool, PGSs, as a measure of "genetic influences" or as being useful for enhancing understanding is not the same thing as denying genetic differences or assuming a blank slate view of human psychology. The critique that I assume psychological homogeneity is both wrong and irrelevant. Indeed, we agree that individuals differ genetically and psychologically in a manner that shapes development and social outcomes. The key question at issue, which Richters avoids, is whether PGSs have utility for enhancing understanding of these differences.

R1.3. Recommendation: Use PGSs sparingly and cautiously given limitations

A few commentators interpreted my argument as being that PGSs are useless and should never be used in social science. For example, in their otherwise concurring response, **Veit & Browning** argue that I overstate my conclusion, which they interpret as being that "sociogenomics is methodologically doomed" and that PGSs are useless for all social science purposes no exceptions. This ostensible disagreement is based on misunderstanding. I specifically recommended that PGSs should be used "sparingly and cautiously" for social science rather than "not at all."

Analogously, although agreeing that my critique is "mostly correct," **Fletcher** takes issue with what he perceives to be my argument that PGSs are worthless and aims to carve out a "clear-eyed middle ground."[2] Concurring with my arguments, Fletcher writes that studies representing PGSs as "genetic propensity" and which are using PGSs as "genetic influences" (vs. environmental ones) are "overstepping" and "a fool's errand." Given that most sociogenomics studies use PGSs in this manner, it follows that we agree that most studies use PGSs inappropriately. However, and purportedly disagreeing with my position, Fletcher suggests that "PGSs can be wrong but useful" "in a limited and focused role in social science research." Contra Fletcher, I did not argue otherwise. The "clear-eyed middle ground" Fletcher aims to carve out was that carved out in my article.

R1.4. Downward causation as a confounder *of PGSs*

A few commentators interpreted my discussion of downward causation as implying that PGSs "only" or "merely" reflect artificial

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

(social) causation (**Trejo & Martschenko**; **Xia & Hill**). For example, Xia & Hill write that I describe "the signal captured by a PGS…on social science traits such as education as being 'artificial.'" This is a misconception with benefits, as it allows them to apply to my arguments another label familiar to critics of behavior genetics: environmental determinism (in this case downward determinism). I am surprised by this interpretation not only because downward causation is but one of several confounders to PGSs that I describe, but also because I explicitly rejected an environmentally determinist approach. In the service of explanation, I employed simplified examples to illustrate the point that because of downward causation "genetic associations for many complex social behaviors are unavoidably environmentally *confounded*" not determined (emphasis added). When I wrote: "As is well known, a person's social traits emerge from a complex interplay of environmental and genetic influences over their lifetime," I meant it.

To be clear, my claim that PGSs capture artificial genetic associations does not imply that PGSs *only* capture artificial genetic associations. We agree that an environmentally determinist approach is untenable.

R1.5. No enthusiasm for the outdated nature versus nurture debate

In another unanticipated response, several commentators (**Trejo & Martschenko**; **Alexander et al.**; **Richters**) charge me with "perpetuating the nature versus nurture debate." Although sympathetic with some of my critiques, Trejo & Martschenko write that I "desire to separate nature versus nurture" and my arguments encourage attempts at such separation. Richters claims that my arguments "renew the charter" of genetic versus environmental separation. In all cases, this critique is asserted but not explained, and as I do not see how this follows from my arguments, I cannot engage directly with their reasoning.

To clarify, my discussion of environmental confounding was not meant to encourage efforts to differentiate genetic versus environmental influences, which we agree is a futile endeavor (see Burt, 2015; Burt & Simons, 2014). On the contrary, by illuminating the fallacy in treating PGSs for complex traits as "genetic influences," I was arguing *against* the interminable effort to separate nature and nurture in its contemporary form with PGSs as "nature." When I wrote that studies using PGSs as genetic influences are "fundamentally and necessarily wedded to an overly simplistic and ultimately misleading (environmentally confounded and biologically implausible) reductionist genes-versus-environments approach," and the problem is not tractable with advanced statistical methods, as **Trejo & Martschenko** agree, I meant that too. We "can no more unbraid genetics and environments [on complex social traits] than we can unbraid history and culture, or climate and landscape, or language and thought" (Feldman & Riskin, 2022).

R1.6. Unknowns and false dilemmas

From my claims that we don't need PGSs to show well-established social patterns (e.g., "to demonstrate that supportive, stimulating parenting is associated with child educational attainment"), **Morris et al.** craft a straw man, perhaps for rhetorical effect. They misrepresent me as holding "that we know the answers to all the important questions already." Obviously, we do not.

In a more reasonable objection, **Morris et al.** write that: "environmental causation is precisely what genetically controlled designs help establish in observational research." I anticipated this response, and I refer the reader to sections 5 and 6 of the target article where I discuss why demonstrating environmental causation is not a strength of PGSs. Briefly, because, as we all agree, PGSs do not control for "all genetic differences" and are environmentally confounded, I noted:

> even if the inclusion of PGSs markedly altered an environmental estimate, because PGSs are significantly environmentally confounded, we cannot say that controlling for "genetics" is the cause of such changes. What is more, we cannot say that environments matter "net of genetics" because PGSs only capture a fraction of the ostensible heritability of social outcomes (see also **Fox**; **Zietsch et al.**).

Disappointingly, **Morris et al.** did not engage with these specific arguments. Instead, they pose a dilemma: Support PGSs or support genetically confounded social science research. Fortunately, this is a false dilemma.

R1.7. Not defending "standard social science model" or social measurement

Richters objects to my argument because, in his view, I do not "highlight precisely the same deficiencies in the social science model [I] seek to defend…"[3] Richters' critique is, however, based on a misunderstanding; my target article is not a defense, much less a "vigorous defense," of social science research. There is no contradiction in addressing the challenges with PGSs for social science and holding that social science research, in general, has many challenges, even deficiencies.

In a similar critique, **Morris et al.** complain that were I dispassionate and focused on scientific accuracy, I "would be as concerned about exaggerating the effects of 'structural disadvantages and cultural influences' as 'obscuring them'." I anticipated this tu quoque, and I point the reader to section 8 where I attempted to dispel such unproductive discussions. Manifestly, my target article was not an overview of "problems with social science" but had a very specific focus on challenges with PGSs.

Focusing on measurement, **Moreau & Wiebels** interpret me as holding that "because they are already well-measured behaviorally, constructs like academic achievement or cognitive aptitudes have little to benefit from the tools of sociogenomics." This is a two-part claim, and both are misguided. First, I did not argue that constructs like cognitive aptitudes or psychosocial traits are well-measured. Indeed, I share their concern about the measurement of social constructs (see, e.g., Burt, 2012, 2020) and agree that we "should refrain from thinking that the measurement of constructs in the behavioral sciences is as good as it can be." Second, my critique of the utility of PGSs in social science is not based on adequacy of social measurement. If anything, my arguments would lend support to the claim that the *inadequacy* of measurement of social constructs poses a challenge to GWASs and PGSs. Although I agree that PGSs will be more useful for medical phenotypes defined by the "presence or absence of biological features," *pace* Moreau & Wiebels, the fact that "behaviorally assessed constructs remain subjective and far from assumption-free" is, in my view, a barrier to genetic analysis not an argument for its utility.

R1.8. Context-dependency: More than an interpretive problem

In the target article, I discussed the context-dependency of PGSs and outlined the implications for complex social traits (see sect.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

5.4). In his commentary, **Fletcher** briefly depicts this significant challenge as being of a narrower problem: That the "*interpretation* [of PGSs] is context-dependent" (emphasis added). Although Fletcher is correct in that the interpretation of PGSs themselves – as the aggregate scores – is context-dependent, this framing of the challenge as an "interpretive" one minimizes the complications. The issue is not merely that the interpretation of a PGS effect is context-dependent in the same way that the interpretation of the label "sick" varies from "good" among a group of high school skateboarders, to "disgusting" among people discussing a ghastly crime, to actually "ill," as traditionally defined. The challenge is much more complicated as contexts can shape which and how – that is, the magnitude and even direction – individual genetic variants matter. This context-dependent variation is missed in PGSs, which are weighted aggregates of the average effect of a tag-SNP in a specific context estimated from disproportionately European-genetic ancestry samples that are frequently not representative of the underlying population (e.g., wealthier and more highly educated; **Curtis**, also Burt & Munafò, 2021).

Consider an analogy. If I create a weighted scale of 100,000 individual characteristics associated with success in football (context) and call it "athletic propensity" (PGS), and then I apply this "athletic propensity" algorithm to different athletic contexts like soccer, tennis, cycling, and rowing, it will surely perform less adequately in predicting success. The lower predictive ability of this "athletic propensity" scale does not indicate that athleticism matters less for soccer or tennis, but rather follows from the fact that these sports (as contexts) differ and with it the nature and salience of various skills and capacities associated with success. Additionally, like an educational-attainment PGS, using an additive, unidimensional scale of "athletic propensity for football" is misleading (see also **Richardson**; **Sarkar**). A variety of traits and combinations thereof facilitate success even within the same context, as even the most cursory comparison of characteristics of football players at different positions would suggest. So too for the skills facilitating educational attainment across contexts and even for different subjects like fine art and music studies compared to sociology and psychology or physics and chemistry. For complex social traits, context is intertwined with almost everything at the phenotypic level; these contingencies are exponentially more complicated at the genetic level.

I reiterate this important point because the context-specificity of PGSs continues to be underappreciated and contributes to misuse (see citations in the target article, **Curtis**; **Moore**; **Sarkar**). In particular, existing studies and claims about the potential utility of PGSs are insufficiently attentive to the implications of the context- and condition-dependent nature of PGSs (but see Mostafavi et al., 2020). To reiterate, I was not arguing that this context-dependency makes PGSs useless. Rather, I was highlighting how this context-dependency undermines their utility for certain usages – for example, comparing PGSs across contexts to assess variation in "how much genetics matters."

R1.9. Mine is a scientific not ethical or sociopolitical critique

Controversies about the ethical and sociopolitical implications of including genetics in social science are longstanding. Distinguishing my target article from extant critiques of sociogenomics, I noted that most existing critical engagement focuses on sociopolitical and ethical concerns.[4] These works address questions such as: Is it ethically responsible to study the genetics of social outcomes profoundly shaped by inequality? How should findings from the field of sociogenomics be used? Who stands to benefit? Who will be harmed (or will not benefit)? And do these ethical concerns about this work outweigh the scientific gains?

These are not the questions addressed in my target article. My article focuses on the scientific challenges with PGSs and the implications for social science, as several commentators recognize (e.g., **Trejo & Martschenko**). Scientific questions I address include, for example: What do PGSs measure? Do PGSs indicate "genetic influences" on complex social traits as they are often used? Given, as I discuss, they do not, what is their scientific utility for enhancing understanding of social behavior?

Nonetheless, some commentators charge me with being motivated by sociopolitical and ethical concerns. **Morris et al.** allege that I conflate scientific and ethical concerns, pointing as evidence to my conclusion that the "*scientific* costs outweigh the meager benefits" (emphasis added).[5] My purported non-scientific concerns about PGSs they cite include "obscuring environmental influences," "perpetuating a flawed concept of genetic potential," and "wasting resources."[6] Given the goal of social science of explaining variation in social behavior, inasmuch as PGSs obscure environmental influences and perpetuate a flawed concept of genetic potential (which is my argument), this impedes scientific advancement (i.e., is a scientific cost). Morris et al. disagree, arguing that the limitations I outline are overstated and/or tractable and thus my recommendation to use PGSs sparingly and cautiously in social science is not justified. However, *our disagreement is scientific* not sociopolitical or ethical.

R2. Assorted genuine disagreements

The previous section outlined ostensible disagreements rooted in misunderstanding. In this section, I address genuine disagreements grounded in disputes about the facts or their implications.

R2.1. PGSs and evolutionary insights

Two commentaries draw upon evolutionary perspectives to critique or refine my arguments. Focusing on the utility of PGSs, **Hong** argues that I overlooked their value for "greatly and uniquely" contributing to "the study of genetic evolution in contemporary societies." In the target article, I necessarily focused on key arguments about the utility of PGSs for social science. In my reading, enhancing understanding of "natural selection in contemporary human populations" is not a common or touted use of PGSs in social science. This is evidenced by the paucity of such studies and the absence of discussion of the utility of PGSs for such purposes in salient overview articles (e.g., Harden & Koellinger, 2020; Mills & Tropf, 2020). Notably, the limitations of PGSs I discuss also impede their utility for the aim of understanding selection in contemporary human populations (e.g., Berg et al., 2019; Sohail et al., 2019).

As noted in the target article, sociogenomics research tends to suffer from a deficit of theory, including evolutionary theory. This manifests in the dearth of theoretically driven models and concepts, including phenotype selection (Boardman & Fletcher, 2021; Burt, 2022, 2023, also **Charney**). Drawing on evolutionary theory, **Ramus** concurs with my argument that complex social traits, like educational attainment, are not well-suited for GWASs and PGSs because, in his view, "these complex social outcomes are not phenotypes that are under direct natural selection." The solution, according to Ramus, is "redirecting geneticists'

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

attention to stable traits [that] can be defined and can be the target of selection." Focusing on cognitive outcomes, he suggests that components underlying specific cognitive abilities, such as verbal ability, working memory, or number sense, as well as character traits, like self-control, intrinsic motivation, and grit are more appropriate phenotypes as they are relatively stable and under direct natural selection.

I concur with **Ramus** that more narrowly defined, stable traits that are the target of direct natural selection are more appropriate traits for genetic analysis than emergent, social achievements like educational attainment. However, I disagree that the cognitive and character traits he identifies meet these criteria – that is, are appropriately viewed as stable traits that are the target of direct natural selection. Scholarship in evolutionary-developmental behavioral science undermines the notion that such cognitive traits are stable or under direct natural selection (as being uniformly fitness promoting). After all, we did not evolve to maximize wealth, educational attainment, happiness, or even health but to survive and reproduce. Moreover, evolutionary-developmental models direct theoretical attention away from the single-"best" traits (e.g., future time orientation, conscientiousness, working memory, task persistence) and toward context- and condition-dependent optimal traits (see, e.g., Belsky, Steinberg, & Draper, 1991; Chisholm, 1999; Ellis et al., 2022; West-Eberhard, 2003). For the intelligence or educational-attainment traits of interest to Ramus, this implies rather than a one-context fits all model of intelligence, an ecologically contingent notion of "adaptive intelligence" (also "successful intelligence" or "multiple intelligences") (Gardner, 2017; Sternberg, 2019).[7]

Regarding stability, evolutionary-developmental models also recognize that contexts are constantly changing and the future is uncertain (Boyce & Ellis, 2005). Given this reality, humans have evolved neurobiological mechanisms facilitating adaptive phenotype plasticity in response to external environmental factors and relative condition (e.g., relative health, status) (Del Giudice, Ellis, & Shirtcliff, 2011). Consistent with this model, a growing body of research over the past two decades demonstrates that rather than being stable, many cognitive and character traits are malleable[8] in response to environmental insults (social and physical) (Burt, Lei, & Simons, 2017; Pepper & Nettle, 2017; Shonkoff & Phillips, 2000) and supports, including interventions, training, and even education (Brinch & Galloway, 2012; Brody et al., 2005; Harrison et al., 2013; Hegelund et al., 2020; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Kautz, Heckman, Diris, Ter Weel, & Borghans, 2014), as **Charney** also notes.

In sum, although **Ramus** and I agree that several "conveniently available" social outcomes are not appropriate phenotypes for genetic analyses, we disagree that the solution is studying specific cognitive abilities assuming these are relatively context-insensitive stable traits under direct natural selection.

R2.2. Downward causation = social causation

A few commentators disagree with my arguments about downward causation creating what I call artificial – as socially produced – genetic associations. Importantly, commentators do not dispute that PGSs unavoidably capture social forces like discrimination. What is debated is whether these forces are properly interpreted as being social (my argument) or as genetic effects (dissenters' argument). Before tackling this debate, I briefly address an extension of my discussion of downward causation by two commentators who agree with my arguments.

In a concurring commentary, **Merchant** helpfully offers a more formal definition of artificial genetic associations produced by downward causation as "any association between genomic variants and a given outcome that is forged through social practices rather than biochemical pathways." In a variation of my argument, both Merchant and **Charney** suggest that downward causation is a form of population stratification (PS, i.e., population substructure produced phenotype stratification) (see sect. 5.1.1). Traditionally defined, PS reflects random allele frequency differences between subgroups that are associated with, but usually irrelevant to, the trait. In contrast, genetic variant – trait associations reflecting downward causation, which I label artificial genetic associations, need not be differentiated by population subgroups and are relevant to the phenotype, because of (at least in part) social not genetic causation. An example is an association between genetic variants associated with height or perceived attractiveness and income or educational attainment, partly rooted in the social tendency to favor more attractive and taller people. Given this, as I note in the target article (and contra **Morris et al.**) existing statistical methods designed to mitigate population stratification confounding (e.g., within-family methods) are not able to correct for artificial genetic associations reflecting downward causation. To be sure, the concepts I introduced would benefit from deeper consideration, further revisions, even new labels, and I hope these issues are addressed in future work.

Dissenting commentaries argue that what I refer to as artificial genetic associations are appropriately viewed as genetic causes that are environmentally mediated ("indirect genetic effects" or "vertical pleiotropy") (**Keller**,[9] **Xia & Hill**, **Trejo & Martschenko**, **Burke**). On their account, again using the example of racial discrimination/colorism, skin pigmentation alleles *cause* skin pigmentation differences, which *cause* differences in racial discrimination experiences, which *cause* disparate outcomes. In their view, this makes the experience of racial discrimination and disparate outcomes caused (in part) by racial discrimination (e.g., higher allostatic load, depression, criminal behavior, income, and educational attainment) caused by and thus "indirect genetic effects" of skin pigmentation alleles.

I disagree. In my view, the cause of racial discrimination based (in part) on skin pigmentation is not genetic variants related to skin pigmentation but *social forces* (racism/colorism) that act "downward" on genetically influenced differences such as skin pigmentation (see, e.g., Burt, 2018; Burt, Simons, & Gibbons, 2012; also **Merchant**). There is no biological pathway upward from skin pigmentation alleles to racial animus, discriminatory treatment, or racial segregation. In my view, this idea that racial discrimination/colorism is caused by skin pigmentation alleles is rooted in a misguided gene-centric worldview where causality is something that only occurs in one direction: Upward from lower-level parts to higher-level entities. In contrast, I adopt an ontologically pluralist view, which recognizes that higher-level emergent phenomena (social structures) can operate causally on lower-level factors (see, e.g., Burt, 2023; Dupré, 2012).[10]

Thus, although commentators and I agree that PGSs unavoidably capture social forces like discrimination, we will have to agree to disagree whether the resulting genetic associations are properly viewed as artificial (social) or genetic influences. **Trejo & Martschenko** submit that we need new language to describe such relationships. I do not disagree. In the meantime, however, we likely agree that we would benefit from employing existing concepts with greater care and accuracy.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

R2.3. (In)tractability of limitations

As noted, commentaries generally concur with the challenges and limitations that I outline for PGSs as well as the need for appropriate interpretation. Disagreement centers on the tractability of the limitations and my recommendation to use PGSs sparingly and cautiously. Several commentators argue that PGS limitations – especially environmental confounding – are adequately mitigated with sophisticated methods and do not undermine the utility of PGSs for social science (e.g., **Keller**; **Morris et al.**; **Zietsch et al.**). As discussed in the target article, although I agree that several of these methods substantially mitigate confounding, I disagree that the environmental confounding of PGSs is an issue that can be overcome with statistical genetic methods.

I shall not repeat my arguments from the target article explaining why methodological limitations and gene–environment interplay in development result in unavoidably environmentally confounded PGSs (see also **Archer & Lavie**; **Charney**; **Curtis**; **Moore**; **Richardson**; **Trejo & Martschenko**). Nor shall I reiterate, in response to **Morris et al.** who suggest these problems are "scientifically tractable issues that have been substantially addressed," the limitations, questionable assumptions, and new challenges accompanying these novel methods (see Zaidi & Mathieson, 2020; also, Boardman & Fletcher, 2021; Charney, 2022; Domingue & Fletcher, 2020; Young et al., 2022). Instead, I point readers to section 5 of the target article, where I discuss these issues along with section R.2.4.

Conversely, several commentators suggest that the problems with PGSs are worse than I outline, for example, arguing that PGSs have "negative utility" and that social scientists should "steer clear" of them (e.g., **Curtis**; **Richardson**; **Sarkar**). Although my position is admittedly closer to the "useless" than the "very useful" arguments of some commentators (e.g., **Alexander et al.**), my arguments fall in the middle. I neither suggest that PGSs should never be used or can never be useful in social science. I argued that their utility is narrow. My position was solidified by advocates' meager examples of utility given limitations.

Rather than tractable, my view and that of many other commentators is that the various difficulties plaguing PGSs as reflecting "genetic influences" for complex social traits are insurmountable. Leaving aside the question of whether disentangling socio-environmental influences from authentic genetic signals influencing complex social traits is possible in principle – and I and many commentators think not – we are nowhere near there yet. Some commentators admit as much but suggest "this is no reason for despair" but rather we should continue to try to develop innovative strategies to overcome these limitations (**Keller**; **Morris et al.**). I neither counsel despair nor oppose the development of innovative strategies, as implied. In fact, I encouraged the use of more robust, innovative strategies (e.g., sibling difference GWASs) in the target article. What I oppose is the misuse of PGSs, as, for example, representing genetic (vs. environmental) influences.

R2.4. Challenges with novel advanced methods

Although scrutinizing various specific methods designed to mitigate confounding in GWASs and PGSs is impracticable here, I nevertheless wish to briefly address **Keller**'s response to my questioning strong assumptions underlying a popular contemporary approach to mitigate confounding in GWASs (LDSC; Bulik-Sullivan et al., 2015) and PGSs creation (e.g., LDPred; Vilhjálmsson et al., 2015). I wrote that in all applications of LDPred that I have seen, studies assume "that all SNPs are causal," which "is curiously not defended anywhere to [my] knowledge," and "not consistent with available empirical evidence." Keller implies that my questioning such assumptions reflects scientific naiveté or unreasonableness, given that "models are not meant to mirror reality."[11] In his view, at issue is "the degree to which results are biased and whether this bias matters."

I agree. This is why much of my target article was focused on explaining the degree and, especially, import of these biases for social outcomes. Although we cannot know the precise amount of bias, given the nature of development and methodological limitations, the evidence we do have suggests it is both substantial (e.g., as little as 1/3 of the EA PGS effect is attributable to "direct genetic effects"; Okbay et al., 2022), insufficiently corrected (e.g., Young et al., 2022; Zaidi & Mathieson, 2020), and matters for understanding (Berg et al., 2019; Haworth et al., 2019; Sohail et al., 2019). To be clear, my point was not about any one of the specific dubious assumptions involved in these various methods, but more broadly on the reliance on assumptions, sometimes strong, at almost every level of analysis that are questionable, rarely justified, often obscured, and frequently unknown by non-experts who apply the products of these techniques in social science applications and present the resulting outcomes as being "corrected for" environmental confounding and other issues.

R2.5. Wrong but useful for prediction?

Several commentators argued that the environmental confounding of PGSs does not undermine their utility for risk prediction (e.g., **Keller**). On this view, PGSs are valuable as they offer practically useful incremental prediction that is independent of traditional social measures (**Alexander et al.**; **Moreau & Wiebels**). That the prediction accuracy of PGSs is not undermined by environmental confounding is, of course, true, with the usual caveats (e.g., context-dependency). However, that PGSs have actionable utility for predicting individual risks for complex social traits at the current state of the science is widely recognized to be misguided (but see Plomin & Von Stumm, 2022). As noted in the target article, most scholars, including ardent supporters of PGSs for social science, "agree that PGSs do not predict complex social outcomes with any degree of efficacy or accuracy and, therefore, should not be used for individual prediction" (citations omitted, see also **Moore**; Turkheimer, 2015, 2019). Moreover, as noted in the target article, research suggests that the incremental predictive efficacy of social science PGSs independent of available or easily attainable phenotypic measures, such as prior grades or parental educational attainment, is too weak to be of practical utility (e.g., Morris, Davies, & Smith, 2020).

Furthermore, we still have the "ancestry-specific" "portability problem" (see sects. 2.3 and 3.1). As **Curtis** elaborated, because PGSs are tailored to specific ancestral populations in certain contexts, different PGSs would have to be created for different ancestries (Martin et al., 2017). Further complicating matters, the human population cannot be neatly demarcated into different ancestral groups. "In reality, there is no bright line demarcating comparisons 'within' versus 'between' ancestries: there is a giant family tree of humanity, and people who share more ancestral paths through it than others, and more similar environments than others" (Coop & Przeworski, 2022, p. 850).

R2.6. Incautious usage

Although equivocating somewhat, some commentators seem to suggest that most social science studies use PGSs appropriately

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

(**Alexander et al.**; **Morris et al.**). Given their position, you might think that most social science studies eschew depictions of PGSs as "genetic influences" and adopt rigorous methods to mitigate confounding. You would be mistaken. I refer the reader to sections 5 and 6 of the target article where I highlight several misguided applications of PGSs as "genetic propensity," "genetic influences," even "genetic endowment of educational attainment." On my reading, most – but certainly not all – current social science applications use PGSs as "genetic influences."

Also noted in the target article, only a paucity of sociogenomics studies employ PGSs using the most robust available methods designed to attenuate environmental confounding (including within-family studies highlighted by **Morris et al.**, **Keller**, **Zietsch et al.**, and Madole & Harden, 2023).[12] Some studies, cited in the target article, fail to employ even basic, necessary adjustments in the creation of PGSs (e.g., LD pruning). In still other cases, studies provide insufficient detail on PGSs necessary for an evaluation, for example, saying only that the PGS was "based on the effect sizes from the most recent GWAS of educational attainment."

In short, incautious application with insufficient description and/or corrections, overinterpretation, and misinterpretation of PGSs as genetic propensity is a significant problem (see also **Fletcher**). Often problematic use is accompanied by overinterpretation, hype of weak evidence, and promissory notes. This brings us to the commentary by **Alexander et al.** whose views starkly diverge from mine.

R2.7. Hype and disparate interpretations

In their commentary, **Alexander et al.** offer a "defence of the immediate practical utility of PGSs for maximizing trait *prediction*" and for "advancing etiological *understanding*" of complex social traits.[13] In my view, their defense epitomizes what I consider to be an incautious hyping of PGSs based on misinterpretation and which my article is intended to counter. However, I am grateful for their response, which encourages direct engagement with opposing claims for the utility of PGSs. Their arguments sound compelling but are, in my view, partial, misguided, or based on questionable assumptions or weak evidence.

To the utility of PGSs for prediction, **Alexander et al.** assert that: "it is only legitimate to assume that PGSs [for complex social traits] are just about to unfold their full predictive potential." For all the reasons I have discussed, I disagree.

Alexander et al.'s defense of the utility of PGSs for etiologic understanding crumbles under scrutiny. Space does not permit a critical citation-by-citation analysis of the support they present for their perspective, so one example will have to suffice. They write:

> The growing number of studies combining PGSs with neuroimaging, proteomic or other multi-omic data have already provided unique insights into *specific mechanisms* through which *polygenic predispositions exert their effects* on complex phenotypes. Exemplary findings from neuroimaging studies include the identification of *structural brain changes* associated with PGSs for neuroticism (Opel et al., 2020) and educational attainment (Elliott et al., 2019), that, in the latter example partly mediated the association between participants' PGS and their cognitive test performance (emphases added).

This evidence is presented to contradict my claim that PGSs lack utility for identifying specific biological pathways to social outcomes. However, **Alexander et al.** mischaracterize these studies. Neither study examines "structural brain *changes*"; both studies analyze brain measurements at a single time point. For example, Elliott et al. (2019) test the hypothesis that an educational-attainment PGS influences individual differences in intelligence by "contributing to the development of larger brains," which could "constitute a biological pathway linking genetic variation to differences in intelligence and educational attainment" (p. 3497). Despite what Alexander et al. imply, the results are not particularly noteworthy. The educational-attainment PGS explained less than half of 1% of the variance in brain size. Not surprisingly, the mediation analyses revealed extremely weak indirect effect sizes ($b = 0.01$), with significant effects observed in only two of the four samples. Moreover, Elliott et al. did not follow recommended protocols to mitigate biases in PGSs as discussed in the target article.[14] Given the paltry effect sizes, one might reasonably expect that these estimates would be naught if such adjustments for confounding had been implemented.

Leaving aside concerns about methodological limitations, these cited studies illustrate my concern that PGSs are being used in a manner that obscures potentially relevant socio-environmental influences (*pace* **Morris et al.**, for scientific not ethical reasons). **Alexander et al.**, following Elliott et al., assume that the educational-attainment PGS *causes* structural brain differences, which *cause* educational-attainment differences. However, this causal ordering cannot be assumed given the significant environmental confounding of PGSs (see also Coop & Przeworski, 2022). Our brains are co-constructed from combined genetic and environmental influences in development. Studies of neuroplasticity in human and rodents demonstrate that a variety of socio-environmental forces alter the structure and function of the brain and with it our ability to respond to ongoing challenges and opportunities (e.g., Glasper & Neigh, 2019; Kokras et al., 2019; Leuner, Glasper, & Gould, 2010; Sweatt, 2016). We all know the life experiences of the average person who gets a Ph.D., J.D., or M.D. and the person who does not graduate high school are very different. These different experiences and contexts of development, which include the experience of education itself, are neurobiologically embodied. For these reasons, interpreting differences in brain structure or cognitive test performances as reflecting causal genetic differences based on PGS associations is unjustified.

In sum, where **Alexander et al.** dispute my arguments and assert that "PGSs hold great potential for both better *prediction* and *understanding* of complex traits in social science" (emphasis in original), I find their evidence problematic and uncompelling, and I strongly, albeit respectfully, disagree.

R3. Miscellaneous agreement and extensions

In this section, I consolidate additional extensions, amplifications, and points for fertile discussion.

R3.1. Downplaying wider social forces

Expanding on a brief critique (see note 10 of the target article), **Merchant** argues that so-called "dynastic effects" (or "indirect genetic effects") are "undertheorized and underexplored" in GWAS/PGS studies and "often assumed to describe the direct genetic effects of the parents' genotypes on their parenting" (see sect. 5.1.2; also Coop & Przeworski, 2022; Young et al., 2022). We agree that the scholarly discourse around this form of environmental confounding, often framed as "genetic nurture," is overly focused on parenting in a manner that obscures the effect of wider socio-cultural forces. A rich body of social science

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

research highlights the manifold ways that social forces beyond parenting – for example, schools, neighborhoods, social networks – influence important life outcomes (see, e.g., Leventhal & Brooks-Gunn, 2000; Sampson, Raudenbush, & Earls, 1997; Simons & Burt, 2011). Although several recent studies have usefully explicated and empirically demonstrated how PGSs capture these broader social forces (Abdellaoui, Dolan, Verweij, & Nivard, 2022; Young, Benonisdottir, Przeworski, & Kong, 2019), a general tendency to narrowly focus on parenting remains.

R3.2. Broader untenable assumptions

Several commentators expand my critique to highlight broader questionable assumptions underlying PGS studies, including those that have long been critiqued. Both **Richardson** and **Archer & Lavie** criticize the oversimplified additivity assumption and overall "gene-centric approach" underlying much sociogenomics work. They emphasize the dynamic responsivity of cells and organisms to their environments and the role of the genome as a resource facilitating such responsivity. Both underscore the need to replace the "gene-centric approach" with a more "biologically realistic one" (Richardson; also Richardson & Jones, 2019).

In his commentary, **Sarkar** writes that arguments that PGSs have sidestepped the invalid assumptions and environmental confounding of prior eras of social science genetics "are not credible"; I concur. Sarkar helpfully points out that social scientists are not alone in their concerns about the use of GWASs and PGSs for complex social outcomes. After noting that 1970s critics of heritability studies "read like a 'Who's Who' of theoretical population and quantitative" geneticists, Sarkar notes that in recent years prominent geneticists have criticized sociogenomics. This echoes Bliss's (2018) observation that although the behavioral scientists she interviewed "spoke highly of social genomics," she found an "almost polar opposite response from mainstream genome scientists" (p. 157). For example, Bliss wrote that a recent past president of the American Society for Human Genetics expressed concern "about the ways in which social genomic researchers were characterizing social phenomena as medically relevant traits" and remarked that "he hardly believed that any serious scientist would take social genomics seriously" (p. 158). This recognition that contemporary genomics scholars have published critiques of oversimplified assumptions and/or expressed concerns about the application of these genomic tools to study complex social traits (e.g., Coop & Przeworski, 2022; Nelson, Pettersson, & Carlborg, 2013; Rosenberg, Edge, Pritchard, & Feldman, 2019) is a rebuttal to the tacit or explicit suggestion that critics of PGSs are invariably naïve and/or politically motivated social scientists.

R3.3. Socio-environmental epigenetics

Moore and **Gooding & Auger** highlight interesting, important research linking social environmental exposures to epigenetic mechanisms regulating gene expression as a crucial aspect of development and a challenge to PGSs. Both commentaries suggest I was remiss to not discuss environmental epigenetics. Despite their centrality to development, plasticity, and individual differences, epigenetic mechanisms, *per se*, are not a challenge to PGSs, which is why I did not discuss. Specifically, to the extent that environmental influences inducing epigenetic marks are uncorrelated with genotype, they do not confound PGS associations. In the same way that the correlation between parental income and child educational attainment is not undermined by the fact that parents devote money and other monetary resources to children differently, the correlation between a PGS and some outcome is not undermined by epigenetic mechanisms differentially regulating gene expression. Conversely, if environmentally induced epigenetic marks are correlated with PGSs, they reflect one or more of the forms of environmental confounding that I discuss – population stratification, familial confounding, and downward causation – and/or genetic influences.

R3.4. Alternative approaches to overcome limitations

Several commentators concur with my main arguments but suggest shifts in approach or additional analyses to overcome limitations of PGSs. Highlighting the lack of consideration of GWASs in the context of development, **Freitag & Kelsey** recommend adopting a "developmental dynamic" approach and the inclusion of "wider age populations" "to gain a holistic understanding of the biology underlying developmental outcomes." I agree that development is insufficiently considered in GWASs and PGSs. However, for all the reasons I outline in my target article, I believe that the limited utility of GWASs and PGSs for complex *social* traits remain with a developmental perspective. Even so, when considering the application of GWASs and PGSs to disease traits while also recognizing that these studies will continue in various fields regardless of what I say about them, I agree with Freitag & Kelsey's recommendation that adopting a developmental approach would be beneficial.

Addressing the limitations of PGSs for identifying causal variants and biological pathways, **Fox** suggests that strategies to identify "rare variants of large effect" might be useful. Insight from such approaches, Fox argues, could provide knowledge on biological pathways and detrimental mutations, which could be used "to repair or counteract the deleterious effects of the mutation." I agree that rare variants approaches, despite their challenges, have utility for biomedical conditions, like congenital deafness and cystic fibrosis, which reflect a biological dysfunction. However, for normally varying *social* outcomes, like educational attainment, income, smoking behavior, and same-sex sex, I do not share Fox's enthusiasm for rare variant approaches. This is because this approach necessarily rests on the assumption that these social traits reflect a biological deficit produced by a rare variant of large effect. However, as discussed in the target article, the assumption that normal variation in complex social outcomes – for example, "only" graduating high school versus graduating college – reflects a biological deficit is unjustified (see, e.g., Burt, 2023).

Concurring with my arguments that PGSs lack utility for providing biological insights, **Nephew, Murgatroyd, Polcari, Santos, & Incollingo Rodriguez (Nephew et al.)** argue that augmenting PGS studies with "functional (transcriptome, methylome, metabolome) and/or multimodal genetic data," can enhance understanding of biological pathways linking genetics and environments to complex traits and knowledge of "the genetics of social phenomena." I agree with Nephew et al. that sociogenomics studies in social science could be enhanced with such functional genetic data and physiological measurements. However, I remain skeptical of the utility of PGSs even with "additional, more functional assessments of genetic context" for the reasons discussed in the target article. Both because PGSs are environmentally confounded and because they aggregate millions of mostly non-causal variants with miniscule effects, PGSs have limited utility for providing mechanistic insights into complex social traits even when combined with

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

functional and multimodal genetic data. Although I appreciate Nephew et al. drawing attention to these data and possibilities, in my view, these neither rescue PGSs from their limitations nor significantly expand their utility for social outcomes at the current state of the science.

R4. Conclusion

> We're told that science self-corrects, but what the candidate-gene literature demonstrates is that it often self-corrects very slowly, and very wastefully…" Munafò (cited in Yong, 2019).

Early in my target article I noted that this new sociogenomics era has filled the void left by the recent demise of the candidate gene era. I acknowledged the laudable implementation of methodological corrections, such as much-needed attention to statistical power and correction for multiple testing. Later in the article, I warned that "we have been here before," with here being "excitement around genetics, limitations in methodology, and substantial unknown biology." I pointed to the spectacular collapse and the widely acknowledged failure of the candidate gene approach to enhance understanding of social behavior as a lesson that we should continue to heed moving forward.

In his commentary, **Keller** objects to my comparison of this "PGS era" with the candidate gene era. He argues that unlike findings from candidate gene studies, which are likely to be "predominated by false positives," "research findings on PGSs are very different." Keller avers that "PGS findings are largely replicable and PGSs estimate true quantities." Although Keller notes, and I agree, that the candidate gene era "laid bare the fallibility of the scientific process," I believe there are several more specific lessons to be drawn relevant to this PGS era in social science that he downplays. Most notable among these lessons are the challenges with incorporating products of advanced genetic technologies and statistical genetic methods into social science fields generally lacking expertise in these areas. Salient assumptions and limitations of PGSs are often unheeded by social scientists who consume, build upon, evaluate, and even conduct these studies. There is evidence that mistakes are being made and overlooked.

In general, the excitement around new genetic measures and tools can foster hasty, incautious, and misguided application by social scientists who lack training in genetics. Ours remain a scientific environment that rewards novelty and exciting findings, and genetic findings are generally more exciting and newsworthy than other social science findings (e.g., Panofsky, 2014). To be sure, all methodological tools can be misused and misrepresented. However, incognizance among most social scientists of the limitations of PGSs, methodological best practices, and biological unknowns combined with the ease of use and encouragement to use these new genomic tools in behavioral research, create a context vulnerable to PGS misuse and misrepresentation and a very real risk of repeating the scientific costs of the candidate gene era, which include, in **Keller**'s words, "a waste of millions of dollars and researcher time."

The aim of my target article was to draw attention to this scientific situation to promote awareness of and a more critical dialogue about the use and utility of PGSs in social science. Against arguments about the great value of PGSs for social science, I argued that there is a need to rein in the hype about their utility for enhancing understanding of social outcomes, to be more cautious and accurate in description, and to use PGSs sparingly, given known limitations. Most commentators agreed, a few disagreed with my conclusions or what they perceived to be the motivation for my arguments. Moving forward, I hope these discussions continue with the aim of promoting better science.

Notes

1. In my reading, most studies are inconsistently careful; that is to say, social scientists use, describe, and interpret PGSs appropriately in some ways but not others. In my view, this situation results not from intentional misrepresentation but because of space constraints and the complexities of these studies, which are usually conducted by scholars trained in social science not genetics.
2. **Fletcher** also calls my target article "dissonant," "imprecise," and "unfocused," "like other commenters." Disappointingly, Fletcher does not specify what, precisely, is "imprecise" or "unfocused" in my article. In my biased view, my discussions of environmental confounding, including population stratification, downward causation, biological uncertainty, and low resolution, were as precise as the current science allows while also being accessible to a broad audience including those not familiar with sociogenomics methods.
3. In several perplexing arguments, which I wish to briefly address but not highlight, **Richters** writes that my "methodological case against the utility of sociogenomics research rests on a self-refuting thesis about the environmental confounding of PGSs with complex social traits." Not only is my critique not "self-refuted" (and Richters provides no explanation for such self-refutation), but also, I provide a wealth of evidence in the target article that demonstrate this environmental confounding, which Richters ignores. Even commentators who disagree with my conclusions recognize that PGSs are environmentally confounded (e.g., **Morris et al.**; **Zietsch et al.**). Further, Richters asks, "on what authority [am I] asserting, matter-of-factly, repeatedly, and without explanation that environmental effects masquerade as 'genetic influences' in PGS studies." I do not appeal to authority; I point to empirical evidence. I direct interested readers to sections 5 and 6 of the target article, where I cover these issues in detail.
4. Readers interested in such ethical and sociopolitical discussions about sociogenomics/behavior genetics, which have starkly different foci than my target article, can see e.g., Callier and Bonham (2015); Duster (2015); Martschenko (2021); Parens, Chapman, and Press (2006); Reiss (2000); Richardson (2015); Roberts (2015); and Sabatello and Juengst (2019).
5. This is my full sentence they quoted in part: "I argue that *leaving ethical concerns aside*, the potential scientific rewards of adding PGSs to social science are greatly overstated and the scientific costs outweigh these meager benefits for most social science applications" (emphasis added). I assume unintentional, but the omission of "leaving ethical concerns aside" and quotation of remainder as evidence of my ethical concerns is a bit misleading.
6. Although I am perplexed that **Morris et al.** deem my concern with "wasting resources" as evidencing sociopolitical concerns, I am even more puzzled by Morris et al.'s response. They write: "a substantial share of the funding for GWAS comes from private and philanthropic sources who disagree with Burt's assessment." I do not see their point. Perhaps Morris et al. believe that only public funding can be a concern? Or perhaps they believe that such private funding would go to sociogenomics or no other research? Or perhaps their statement that "private funders disagree with [my] assessment" is presented as evidence that I am wrong (a peculiar fallacy of authority argument)?
7. This is, of course, does not imply that differences in such cognitive traits are *only* shaped by environmental variation.
8. Malleable within limits, not infinitely malleable. As before, recognition of malleability does not imply that genetic influences are irrelevant.
9. Notably, **Keller** employs examples that I agree represent indirect genetic effects rather than social causation (e.g., between skin pigmentation alleles, skin pigmentation, and vitamin D levels and/or propensity to skin cancer).
10. Alternatively, or additionally, our disagreement may reflect different understandings of causality and mediation. **Keller**, **Xia & Hill**, and **Trejo & Martschenko** appear to endorse a counterfactual variable substitution effects model of causality (see Madole & Harden, 2023), with causal mediation being demonstrated by *statistical* mediation. A variable (here racial discrimination) is said to statistically "mediate" all or part of the effect size of a causal variable (skin pigmentation alleles) on an outcome if it reduces ("explains")

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

the causal variables' effect on the outcome when controlled. For example, one might introduce a measure of colorism or racial discrimination in a model linking skin pigmentation alleles to, say, educational attainment or depression. If the estimated effect of skin pigmentation alleles is reduced (as we all agree it would be in this example), this would constitute statistical mediation. However, causality is not demonstrated by statistical mediation, which can be observed in the absence of causal mediation if causal ordering is not correctly specified (or all factors are not accounted for). Thus, we can agree that in a statistical model racial discrimination will likely statistically mediate a portion of the effect of skin pigmentation alleles on depression or educational attainment; however, such a finding does not demonstrate that skin pigmentation alleles *cause* racial discrimination. Moreover, the argument that skin pigmentation alleles → skin pigmentation → racial discrimination → lower educational attainment is even less compelling inasmuch as educational attainment is employed as a "proxy for intelligence" or "cognitive ability."

11. Consistent with my critique, the creators of LDPred wrote: "An arguably more reasonable prior for the effect sizes is a non-infinitesimal model, where only a fraction of the markers are causal" (Vilhjálmsson et al., 2015). Unfortunately, the non-infinitesimal LDPred model "is particularly sensitive to model misspecification when applied to summary statistics with large sample sizes…it is also unstable in long-range LD regions" (Privé, Arbel, & Vilhjálmsson, 2020). A revised version of LDPred, LDPred2, has been developed to address some of these issues, but it too necessarily rests on a variety of assumptions (Privé et al., 2020). I have not seen LDPred2 applied in any PGS studies, and I am not yet familiar with its revisions.

12. Few but not zero. Some studies laudably employ more rigorous within-family methods in social science applications (see, e.g., Belsky et al., 2018; Kweon et al., 2020).

13. Throughout Alexander et al. italicize prediction and understanding.

14. For example, they used all available SNPs from an unrelated GWAS and did not clump or prune SNPs for LD (to avoid inflating the effects of SNPs associated with the variant(s) driving the association).

References

Abdellaoui, A., Dolan, C. V., Verweij, K. J. H., & Nivard, M. G. (2022). Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature Genetics*, *54*(9), 1345–1354. doi:10.1038/s41588-022-01158-0

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., … Herd, P. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, *115*(31), E7275–E7284.

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, *62*(4), 647–670.

Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., … Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *eLife*, *8*, e39725. doi: 10.7554/elife.39725

Bliss, C. (2018). *Social by nature: The promise and peril of sociogenomics*. Stanford University Press.

Boardman, J. D., & Fletcher, J. M. (2021). Evaluating the continued integration of genetics into medical sociology. *Journal of Health and Social Behavior*, *62*(3), 404–418.

Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, *17*(2), 271–301.

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences*, *109*(2), 425–430.

Brody, G. H., McBride Murry, V., McNair, L., Chen, Y. F., Gibbons, F. X., Gerrard, M., & Ashby Wills, T. (2005). Linking changes in parenting to parent–child relationship quality and youth self-control: The Strong African American Families Program. *Journal of Research on Adolescence*, *15*(1), 47–69.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., …Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*(3), 291–295. doi: 10.1038/ng.3211

Burt, C. H. (2012). Gottfredson and Hirschi's conceptualization of self-control: A critique of measures and proposals for change. Retrieved from https://www.researchgate.net/publication/278679704

Burt, C. H. (2015). Heritability studies: Methodological flaws, invalidated dogmas, and changing paradigms. *Advances in Medical Sociology (Genetics, Health and Society)*, *16*, 1–44.

Burt, C. H. (2018). Racial discrimination and cultural adaptations: An evolutionary developmental approach. *Advances in Criminological Theory: Building a Black Criminology*, *24*, 207–252.

Burt, C. H. (2020). Self-control and crime: Beyond Gottfredson & Hirschi's theory. *Annual Review of Criminology*, *3*(1), 43–73. doi:10.1146/annurev-criminol-011419-041344

Burt, C. H. (2022). Cultural evolutionary theory is not enough: Ambiguous culture, neglect of structure, and the absence of theory in behavior genetics. *Behavioral and Brain Sciences*, *45*, e157.

Burt, C. H. (2023). Irreducibly social: Why biocriminology's ontoepistemology is incompatible with the social reality of crime. *Theoretical Criminology*, *27*(1), 85–104.

Burt, C. H., Lei, M. K., & Simons, R. L. (2017). Racial discrimination, racial socialization, and crime over time: A social schematic theory model. *Criminology; An Interdisciplinary Journal*, *55*(4), 938–979.

Burt, C. H., & Munafò, M. (2021). Has GWAS lost its status as a paragon of open science? *PLoS Biology*, *19*(5), e3001242.

Burt, C. H., & Simons, R. L. (2014). Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. *Criminology; An Interdisciplinary Journal*, *52*(2), 223–262.

Burt, C. H., Simons, R. L., & Gibbons, F. X. (2012). Racial discrimination, ethnic-racial socialization, and crime: A micro-sociological model of risk and resilience. *American Sociological Review*, *77*(4), 648–677.

Callier, S. L., & Bonham, V. L. (2015). Taking a stand: The genetics community's responsibility for intelligence research. *Hastings Center Report*, *45*(S1), S54–S58.

Charney, E. (2022). The "golden age" of behavior genetics? *Perspectives on Psychological Science*, *17*(4), 1188–1210.

Chisholm, J. S. (1999). *Death, hope and sex: Steps to an evolutionary ecology of mind and morality*. Cambridge University Press.

Coop, G., & Przeworski, M. (2022). Lottery, luck, or legacy. A review of "The Genetic Lottery: Why DNA matters for social equality" *Evolution*, *76*(4), 846–853.

Del Giudice, M., Ellis, B. J., & Shirtcliff, E. A. (2011). The adaptive calibration model of stress responsivity. *Neuroscience & Biobehavioral Reviews*, *35*(7), 1562–1592.

Domingue, B. W., & Fletcher, J. (2020). Separating measured genetic and environmental effects: Evidence linking parental genotype and adopted child outcomes. *Behavior Genetics*, *50*(5), 301–309.

Dupré, J. (2012). *Processes of life: Essays in the philosophy of biology*. Oxford University Press.

Duster, T. (2015). A post-genomic surprise. The molecular reinscription of race in science, law and medicine. *The British Journal of Sociology*, *66*(1), 1–27.

Elliott, M. L., Belsky, D. W., Anderson, K., Corcoran, D. L., Ge, T., Knodt, A., … Ireland, D. (2019). A polygenic score for higher educational attainment is associated with larger brains. *Cerebral Cortex*, *29*(8), 3496–3504.

Ellis, B. J., Abrams, L. S., Masten, A. S., Sternberg, R. J., Tottenham, N., & Frankenhuis, W. E. (2022). Hidden talents in harsh environments. *Development and Psychopathology*, *34*(1), 95–113.

Feldman, M. W., & Riskin, J. (2022). Why biology is not destiny. *The New York Review of Books*, *69*(7). Retrieved from https://www.nybooks.com/articles/2022/04/21/why-biology-is-not-destiny-genetic-lottery-kathryn-harden/

Gardner, H. (2017). Taking a multiple intelligences (MI) perspective. *Behavioral and Brain Sciences*, *40*, e203.

Glasper, E. R., & Neigh, G. N. (2019). Experience-dependent neuroplasticity across the lifespan: From risk to resilience. *Frontiers in Behavioral Neuroscience*, *12*, 335.

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press.

Harden, K. P., & Koellinger, P. D. (2020). Using genetics for social science. *Nature Human Behaviour*, *4*(6), 567–576.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*(12), 2409–2419.

Haworth, S., Mitchell, R., Corbin, L., Wade, K. H., Dudding, T., Budu-Aggrey, A., … Smith, G. D. (2019). Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications*, *10*(1), 1–9.

Hegelund, E. R., Grønkjær, M., Osler, M., Dammeyer, J., Flensborg-Madsen, T., & Mortensen, E. L. (2020). The influence of educational attainment on intelligence. *Intelligence*, *78*, 101419.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*(19), 6829–6833.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (Working Paper: 20749). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w20749.

Kokras, N., Sotiropoulos, I., Besinis, D., Tzouveka, E., Almeida, O., Sousa, N., & Dalla, C. (2019). Neuroplasticity-related correlates of environmental enrichment combined with physical activity differ between the sexes. *European Neuropsychopharmacology*, *29*(1), 1–15.

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press

Kweon, H., Burik, C., Karlsson Linnér, R., De Vlaming, R., Okbay, A., Martschenko, D., … Koellinger, P. (2020). *Genetic fortune: Winning or losing education, income, and health* (Tinbergen Institute Discussion Paper 2020-053/V). Retrieved from http://dx.doi.org/10.2139/ssrn.3682041

Leuner, B., Glasper, E. R., & Gould, E. (2010). Parenting and plasticity. *Trends in Neurosciences*, *33*(10), 465–473.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, *126*(2), 309.

Madole, J. W., & Harden, K. P. (2023). Building causal knowledge in behavior genetics. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X22000681

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., … Kenny, E. E. (2017). Human demographic history impacts genetic risk prediction across diverse populations. *The American Journal of Human Genetics*, *100*(4), 635–649.

Martschenko, D. O. (2021). "The elephant in the room": Social responsibility in the production of sociogenomics research. *BioSocieties*, *17*, 713–731.

Mills, M. C., & Tropf, F. C. (2020). Sociology, genetics, and the coming of age of sociogenomics. *Annual Review of Sociology*, *46*, 553–581.

Morris, T. T., Davies, N. M., & Smith, G. D. (2020). Can education be personalised using pupils' genetic data? *eLife*, *9*, e49962.

Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J. K., & Przeworski, M. (2020). Variable prediction accuracy of polygenic scores within an ancestry group. *eLife*, *9*, e48376. doi: 10.7554/elife.48376

Nelson, R. M., Pettersson, M. E., & Carlborg, Ö (2013). A century after Fisher: Time for a new paradigm in quantitative genetics. *Trends in Genetics*, *29*(12), 669–676.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., … Gjorgjieva, T. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449.

Opel, N., Amare, A. T., Redlich, R., Repple, J., Kaehler, C., Grotegerd, D., … Dannlowski, U. (2020). Cortical surface area alterations shaped by genetic load for neuroticism. *Molecular Psychiatry*, *25*(12), 3422–3431. doi: 10.1038/s41380-018-0236-9

Panofsky, A. (2014). *Misbehaving science: Controversy and the development of behavior genetics*. University of Chicago Press.

Parens, E., Chapman, A. R., & Press, N. (2006). *Wrestling with behavioral genetics: Science, ethics, and public conversation*. JHU Press.

Pepper, G. V., & Nettle, D. (2017). The behavioural constellation of deprivation: Causes and consequences. *Behavioral and Brain Sciences*, *40*, e314.

Plomin, R., & Von Stumm, S. (2022). Polygenic scores: Prediction versus explanation. *Molecular Psychiatry*, *27*(1), 49–52.

Privé, F., Arbel, J., & Vilhjálmsson, B. J. (2020). LDpred2: Better, faster, stronger. *Bioinformatics (Oxford, England)*, *36*(22–23), 5424–5431.

Reiss, M. J. (2000). The ethics of genetic research of intelligence. *Bioethics*, *14*(1), 1–15.

Richardson, K., & Jones, M. C. (2019). Why genome-wide associations with cognitive ability measures are probably spurious. *New Ideas in Psychology*, *55*, 35–41. doi:10.1016/j.newideapsych.2019.04.005

Richardson, S. S. (2015). The trustworthiness deficit in postgenomic research on human intelligence. *Hastings Center Report*, *45*(S1), S15–S20.

Roberts, D. (2015). Can research on the genetics of intelligence be "socially neutral"? *Hastings Center Report*, *45*(S1), S50–S53.

Rosenberg, N. A., Edge, M. D., Pritchard, J. K., & Feldman, M. W. (2019). Interpreting polygenic scores, polygenic adaptation, and human phenotypic differences. *Evolution, Medicine, and Public Health*, *2019*(1), 26–34.

Sabatello, M., & Juengst, E. (2019). Genomic essentialism: Its provenance and trajectory as an anticipatory ethical concern. *Hastings Center Report*, *49*, S10–S18.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science (New York, N.Y.)*, *277*(5328), 918–924.

Shackel, N. (2005). The vacuity of postmodernist methodology. *Metaphilosophy*, *36*(3), 295–320.

Shonkoff, J. P., & Phillips, D. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academy Press.

Simons, R. L., & Burt, C. H. (2011). Learning to be bad: Adverse social conditions, social schemas, and crime. *Criminology; An Interdisciplinary Journal*, *49*(2), 553–598.

Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., … Sunyaev, S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife*, *8*, e39702. doi: 10.7554/elife.39702

Sternberg, R. J. (2019). A theory of adaptive intelligence and its relation to general intelligence. *Journal of Intelligence*, *7*(4), 23.

Sweatt, J. D. (2016). Neural plasticity and behavior – Sixty years of conceptual advances. *Journal of Neurochemistry*, *139*, 179–199.

Turkheimer, E. (2015). Genetic prediction. *Hastings Center Report*, *45*(S1), S32–S38.

Turkheimer, E. (2019). Against decile analysis. Retrieved from https://www.geneticshumanagency.org/gha/against-decile-analysis/

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., … Zheng, W. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics*, *97*(4), 576–592. doi:10.1016/j.ajhg.2015.09.001

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Yong, E. (2019). A waste of 1,000 research papers. *The Atlantic*. Retrieved from https://www.theatlantic.com/science/archive/2019/05/waste-1000-studies/589684/

Young, A. I., Benonisdottir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype–phenotype associations in humans. *Science (New York, N.Y.)*, *365*(6460), 1396–1400. doi:10.1126/science.aax3710

Young, A. I., Nehzati, S. M., Benonisdottir, S., Okbay, A., Jayashankar, H., Lee, C., … Kong, A. (2022). Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nature Genetics*, *54*(6), 897–905.

Zaidi, A. A., & Mathieson, I. (2020). Demographic history mediates the effect of stratification on polygenic scores. *eLife*, *9*, e61548. doi: 10.7554/elife.61548

https://doi.org/10.1017/S0140525X22001145 Published online by Cambridge University Press