Forthcoming Target Article in *Behavioral and Brain Sciences*

Title: Challenging the Utility of Polygenic Scores for Social Science: Environmental Confounding, Downward Causation, and Unknown Biology

Author Information: Callie H. Burt, *Department of Criminal Justice & Criminology; Center for Research on Interpersonal Violence (CRIV), Georgia State University, Atlanta, Georgia, USA 30301;* email: cburt@gsu.edu; homepage: www.callieburt.org

Short Abstract
Pointing to ubiquitous heritability, sociogenomics enthusiasts argue that social scientists should add genetics to their research. Here, I challenge arguments about value of polygenic scores (PGSs) for social science. I explain how inevitable environmental confounding for complex social traits undermines the raison d'être of PGSs of capturing genetic versus environmental influences. I illuminate the sources of persistent, unavoidable environmental confounding in PGSs as well as the unknown biology. 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.

Long 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 due to 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. I argue that leaving sociopolitical and ethical concerns aside, 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.

Keywords: Behavior genetics; Environmental confounding; Gene-environment correlation; Genetic heterogeneity; GWAS; Human potential; Polygenic scores; Population stratification; Sociogenomics; Statistical genetics

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 of 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 varied 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, GWASs 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 x environment era, which was, by and large, a spectacular failure due to methodological limitations and an oversimplified biology (see Charney, 2021; Dick et al., 2015). Suggesting the candidate gene-era "should be a cautionary tale," psychiatric geneticist 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, 2021). With adjustments for multiple testing, attention to statistical power and large samples, and emphasis on replication, among other revisions, this nascent sociogenomic approach has addressed several methodological limitations plaguing the candidate gene approach. As a result, sociogenomic 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 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 et al., 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 due, some proponents have argued, not to valid scientific reasons but to 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 sociogenomic 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 sociogenomic 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 due to 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 sociogenomic findings to support genetic determinist and inferiorizing claims (e.g., Bliss, 2018; Duster, 2015; Harden, 2021a; Herd et al., 2021; Martschenko et al., 2019). While 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 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 section 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 GWASs associations and PGSs.[1] I then

[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, for example, the issue of selectivity (non-generalizability) of samples that predominant in GWAS (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."

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 are available in several accessible social science data sets 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

****Table 1 (Acronyms & Definitions) here****

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 50bp 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 50bp to more than a million base pairs (1Mbp). 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 detail).

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 to 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. While 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 African 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 structural variants have a disproportionate role in shaping human differences compared to other variants (Chiang et al., 2017; Collins et al., 2020; Takumi & Tamada, 2018). Structural variants can involve multiple copies of genes or the deletion of a gene and thus influence gene dosage (the production of too much or too little of the functional protein). Sudmant et al. (2015) estimated that structural variants were 50 times more likely than SNVs to affect gene expression and three times more likely to be associated with a trait difference than a 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 U.K., Backman et al. (2021) observed an average of ~ 600 variants, including 50 putative loss-of-function (pLOFs)

[2] Or 4-5 nucleotide differences every 1000 bp accounting for structural variants.

variants, per gene. Backman et al. (2021) estimated that on average each of us carries 214 putative loss-of-function variants as 'defective' gene copies. Although this variation is non-trivial, 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 make more polypeptides from the same mRNA transcript) (see Strachan & Reed 2018).

In addition to the several million genetic variants passed down by each of our parents, we inherit roughly 30 to 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 & Continental Populations

Most sociogenomic 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., 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 100k years (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 et al., 2003).

The vast majority of variants in an individual's genome are shared by all continental populations (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 due to 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 frequencies across difference groups poses a major challenge for GWAS by inducing or inflating genetic associations through confounding between genotypes and outcomes (e.g., Berg et al., 2019; Morris et al., 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

[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).

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 1-2kb regions, known as recombination hotspots, which occur every 50-100kb across the genome (Myers et al., 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 represent ancestral segments defined by common, ancient SNPs. Rarer variation exists as heterogeneity around the common, ancient SNPs that define haplotype blocks (Strachan & Read, 2018).

This haplotype structure of our genomes undergirds the GWAS methodology. Measuring and testing each of our 3bn 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 million of these preselected SNPs, known as 'tag SNPs' and a trait. Significant SNPs 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 thousand to more than 1Mbp.

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, CNVs, and structural variants—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 GWAS 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 to 2 million SNPs, and most contemporary GWASs now include ~10 million SNPs, most imputed (Tam et al., 2019).

---

[4] For example, 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.

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 (>1000) GWAS hits and miniscule effect sizes (see e.g., Backman et al., 2021; Crouch & Bodmer, 2020; Edwards et al., 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 GWAS, 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 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 are 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 biallelic 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:

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

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 et al., 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 GWAS 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 GWAS 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 et al., 2020; Schaid et al., 2018). Due to 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 $r^2>.1$[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.

[6] As we have noted elsewhere (Burt & Munafò 2020), these various thresholds are somewhat arbitrary and vary across studies, increasing, as others have also noted, researcher degrees of freedom (Charney 2021).

As noted, risk loci range in size from ~50kbp to over 1Mbp (e.g., Lee et al. 2018). Thus, a GWAS that reports 1237 lead SNPs can thus be understood as identifying 1237 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 structural variants 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 .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 versus 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 (PRS) or genetic risk scores (GRS), 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 et al., 2009; Wray et al., 2007). PGSs operate under a massively polygenic, additive model (Boyle et al., 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' in selective plant and animal breeding in human populations (Meuwissen et al., 2001), 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 et al. 2020, p. 465) or a "single quantitative measure of genetic predisposition"(Mills et al., 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 detail), 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. There are several researcher decisions involved in PGS construction worth noting. 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 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).

---

[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 due to 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.

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 et al., 2020b). Although not appropriate for predicting individual outcomes, proponents emphasize myriad 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., 2018), (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 (GxE) with "robust measures of genotype", in contrast to the limited candidate GxE 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 x 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 due to poor individual 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 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 et al., 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, 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 due to (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 & 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 due to non-random mating. Population structure arises due to 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).

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. While there may be some genetic variants affecting our ability to handle chopsticks (e.g., finger dexterity), most genetic associations would be due to 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 polygenic scores, which reflects sociocultural and physical environmental influences (Abdellaoui et al., 2021; 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 U.K., Finland, the Netherlands, and Western France (e.g., Bycroft et

al., 2019; Byrne et al., 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. (2021) demonstrate that controlling for geographic region decreases heritability signals for 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 ~2500 years) genetic structure with sharp effects poses a particular problem for GWAS/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., 2021; 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., 2021; Lawson et al., 2020). Crucially, these biases are exacerbated under the very modeling conditions most often utilized for social science outcomes – when multiple studies are meta-analyzed and millions of SNPs are aggregated in polygenic scores. In these situations, 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 due to uncorrected population substructure (Abdellaoui et al., 2021; Curtis, 2018; Lawson et al., 2020).

5.1.2 Familial Confounding[10]

Biological parents not only pass on ½ 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

[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'Conner 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., 2021).

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 sociogenomic studies have illuminated the extent of familial confounding in PGSs. These studies suggest that roughly half of the effect of the education PGS is due to familial confounding. For example, Kong et al. (2018) found that controlling for an education PGSs created from parents' non-transmitted alleles (i.e., the other ½ 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=.074 versus .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 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 ~22k 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 GWAS. 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. 2020).

Not unexpectedly, sibling-differences 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., 2021; 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 sociogenomic findings. For example, Howe et al. (2021) 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., 2021), or, preferably, (b) use sibling-study adjusted PGSs through a two-stage model to reduce

[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).

(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. Due to 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 socio-cultural 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 & Artificial Genetic Signals

Downward causation—defined as socio-cultural 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]

---

[12] Importantly, although sibling difference studies significantly reduce environmental confounding, they do not eliminate it; as Zaidi and Mathieson explain, although estimates are unbiased, stratification in the PGSs persistent 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.

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 differences in educational attainment in this hypothetical, in our 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, we presume) make hair color unrelated to educational attainment (and any remove any red-hair genetic associations with education). While 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 Jr et al., 2021; Simons et al., 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 socio-culturally 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'). Due to 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 structural variants make GWASs possible, but it also makes PGSs uncomprehensive 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, due to pervasive gene-gene interactions and interactions between non-coding RNA genes and coding genes, focusing on additive effects from

[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 et al., 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).

tag SNPs is necessarily misleading (as oversimplified) about the true nature of genetic influences (Belsky & Israel, 2014; Zuk et al., 2012). Almost everything that happens even at the cellular level is due to 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 & 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 USA, 2010 India, 2002 Nigeria, 1950 Thailand, or 2021 USA are likely to be distinct in non-trivial ways. Whereas a woman going to college in the USA in 2020 would be conforming, a woman going to college in 1870s USA 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, 2022).

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., 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.

[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, 2022; also Dupré 2012; Lewontin, Rose, & Kamin 1984; Richardson 2017).

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 socio-cultural 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 context or condition. Traits that facilitate educational attainment, and any genetic contributions thereto, are dependent on socio-cultural influences. For example, if physical education classes were equally emphasized with non-PE courses and graded not by effort but 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 context 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 20th century, the potential for educational attainment for women in the USA 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 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 et al., 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—i.e., 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). While 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

[18] Although ethical considerations are not our focus, I question the notion of targeting interventions to those who might need extra support due to high genetic risk versus 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 (see, e.g., Matthews et al. 2021).

groups with different social behaviors and attainments shaped by prior and existing unequal arrangements as 'genetic potential' and, then, excusing future patterns as inevitable due to genetic propensities, even for traits that are substantially driven by social inequalities and malleable.

These studies are in no way unique among sociogenomic 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 to identify 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 due to 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, Forthcoming; 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 sociogenomic 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, that 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.

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 PGS 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,

[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).

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 for 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—i.e., 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 due to population stratification, familial confounding, and downward (socio-environmental) causation. Although methods exist to mitigate the former, especially sibling difference 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 context and condition 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 (versus 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 sociogenomic 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 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.

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 due to 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. Due to these limitations, PGSs cannot be employed as measures of 'genetic influences' as they are being utilized 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.

Funding Statement: Support for this research was provided by a 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.

Conflict of Interest: None

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.

[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.

References

Abdellaoui, A., Verweij, K. J., & Nivard, M. G. (2021). Geographic confounding in genome-wide association studies. *bioRxiv*.

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*. 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). 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). Irreducibly social: Why biocriminology's ontoepistemology is incompatible with the social reality of crime. *Theoretical criminology*, 13624806211073695.

Burt, C. H. (Forthcoming). Cultural Evolutionary Theory is Not Enough: Vague Culture, Neglect of Structure, and the Absence of Theory in Behavior Genetics. *Behavioral and Brain Sciences*

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. (2021). *The 'Golden Age' of Behavior Genetics*.
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., Domingue, B. W., Cesarini, D., Dawes, C., Rietveld, C. A., & Boardman, J. D. (2015). Is the effect of parental education on offspring biased or moderated by genotype? *Sociological Science, 2*, 82.
Conley, D., & Fletcher, J. (2017). *The genome factor*: Princeton University Press.
Coop, G., & Przeworski, M. (2022). Lottery, luck, or legacy? *Evolution*.
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.
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.
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.
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).
Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74.
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*, 00221465211018682.
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. (2021). Within-sibship GWAS improve estimates of direct genetic effects. *bioRxiv*.
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 New York.
Loos, R. J. F. (2020). 15 years of genome-wide association studies and no signs of slowing down. *Nature communications, 11*(1). doi:10.1038/s41467-020-19653-5
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*.
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*. 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.
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
Panofsky, A. (2014). *Misbehaving Science: Controversy and the Development of Behavior Genetics*: University of Chicago Press.
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
Pluess, M., & Belsky, J. (2013). Vantage sensitivity: individual differences in response to positive experiences. *Psychological bulletin, 139*(4), 901.
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*.
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

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, 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*. doi:10.7554/elife.39702
Strachan, T., & Read, A. P. (2018). Human molecular genetics.
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
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).
Yong, E. (2019, May 17). 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*. 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.

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 a 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	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 50bp (includes short CNVs).
Linkage Disequilibrium	LD	The non-random association of alleles on a chromosome; when 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. Region can vary from > a 1Mb to a single base position.
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 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 occasionally 3) 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 >1kb, but with next generation sequencing, SVs it has come to represent smaller changes).
Tag SNPs		Mostly non-functional SNPs in GWAS 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 & Reed 2018).

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.[1] 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 1 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 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 utilized 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 2 or more codons, due primarily to 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 about ~20,000 (protein-coding) genes[2], 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.

---

[1] 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.

[2] The number of human genes is continually updated (revised up and down) and varies across official counts due to slight differences in definitions of genes but has stabilized around 20,000. The number can never be an exact one given variation.

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 et al., 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*)*. A 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'.[3] 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 due to how one defines 'common' (Auton & al., 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 GWAS and used in the creation of PGSs.

A.2.1.2 (Short) Insertion-Deletions (Indels)

A second class of variants is comprised of short insertions and deletions (indels), which includes duplications, deletions, or insertions up to 50bp. (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), are DNA rearrangements (deletions, duplications, or inversions) involving more than 50bp. In the past SVs were defined as larger sequence changes typically up to 1kb, but now are defined as smaller changes and include copy number variants larger than 50bp (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 2x more) overall nucleotide (sequence) differences

[3] Geneticists are moving away from the SNP-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 (<.5%) (Strachan & Read 2018). However, given that the GWAS field uses the term SNP, I will do so here.

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 et al., 2008). Long-read sequencing suggests that there may be several-fold more SVs that are hidden due to systematic biases in detection (Sudmant et al., 2015).

A.2 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 versus 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) to → GGA (glycine)), akin to just adding more ingredients to a recipe.

Unlike SNVs, indels and structural variants affect more than one 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 sequence 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.3 ***A Brief Elaboration of Downward Causation***.

As, I discuss in the main text, the counterfactual 'variant substitution effect' model underlying GWASs and PGSs cannot distinguish between authentic 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 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[4], 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.

References

Auton, A., & al., e. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

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.

Doolittle, W. F. (2013). Is junk DNA bunk? A critique of ENCODE. *Proceedings of the National Academy of Sciences, 110*(14), 5294-5300.

ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia of DNA elements) project. *Science, 306*(5696), 636-640.

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

Hu, X.-S., Yeh, F. C., Hu, Y., Deng, L.-T., Ennos, R. A., & Chen, X. (2017). High mutation rates explain low population genetic divergence at copy-number-variable loci in Homo sapiens. *Scientific reports, 7*(1), 1-13.

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.

Pennisi, E. (2012). ENCODE project writes eulogy for junk DNA. In: American Association for the Advancement of Science.

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.

Strachan, T., & Read, A. (2018). Human Molecular Genetics. Garland. *New York*.

Strachan, T., & Read, A. P. (2018). Human molecular genetics.

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

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

[4] This is basic illustration illustrating 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.