

# Reconstructing Sociogenomics Research: Dismantling Biological Race and Genetic Essentialism Narratives

Pamela Herd<sup>1</sup>, Melinda C. Mills<sup>2</sup>, and Jennifer Beam Dowd<sup>2</sup>

## Abstract

We detail the implications of sociogenomics for social determinants research. We focus on education and race because of how early twentieth-century scientific eugenic thinking facilitated a range of racist and eugenic policies, most of which helped justify and pattern racial and educational morbidity and mortality disparities that remain today, and are central to sociological research. Consequently, we detail the implications of sociogenomics research by unpacking key controversies and opportunities in sociogenomics as they pertain to the understanding of racial and educational inequalities. We clarify why race is not a valid biological or genetic construct, the ways that environments powerfully shape genetic influence, and risks linked to this field of research. We argue that sociologists can usefully engage in genetics research, a domain dominated by psychologists and behaviorists who, given their focus on individuals, have mostly not examined the role of history and social structure in shaping genetic influence.

## Keywords

education, gene-environment interactions, race, sociogenomics

We examine the implications of the recent surge in sociogenomics research for the sociology of health. Relationships between educational attainment, race, or more specifically racism, and health have been central to the sociology of health. Given how early twentieth-century eugenic thinking was used to justify a range of racist and eugenic policies and practices that shape contemporary racial and educational inequality, it is not surprising that many sociologists have serious concerns that sociogenomics research can lead to reductionist understandings of racial and educational inequalities (Galton 1869; Leonard 2017). Consequently, we unpack key controversies and opportunities in sociogenomics as they pertain to the understanding of racial and educational inequities.

We first detail why race is not a relevant biological variable in genetic research, in particular by describing the ancestry concept in the field. We include sociological research that challenged and

influenced how geneticists conceptualized race. We also outline the consequences of excluding non-European ancestry groups from genetics research. Finally, we chronicle ongoing risks regarding how genetics research is presented and misinterpreted, which reinforces false biological narratives of race and distorts the fact that racism is the driving cause of racial disparities in health (Williams, Lawrence, and Davis 2019).

We further scrutinize evolving research on gene-environment interplay in shaping educational attainment. We detail existing evidence, including

<sup>1</sup>Georgetown University, Washington DC, USA

<sup>2</sup>University of Oxford, Oxford, UK

## Corresponding Author:

Pamela Herd, Georgetown University, McCourt School of Public Policy, 37th and O Street NW, Old North #100, Washington, DC, 20057, USA.

Email: [ph627@georgetown.edu](mailto:ph627@georgetown.edu)

the development of a polygenic risk score for educational attainment, with a particular focus on evidence of gene–environment interplay. We also outline misrepresentations of this research that reinforce a narrative of genetic determinism harkening back to early twentieth-century eugenics arguments. We argue that sociologists should engage in this research given our unique theoretical, substantive, and methodological tools, which emphasize history and social structure in shaping genetic influence, thereby helping to ensure that genetics research is not reductionist.

## RACISM AND SOCIOGENOMICS RESEARCH

A key concern raised by sociologists regarding sociogenomics is the history of racism in genetics and the use of genetic arguments to justify racism. As Dorothy Roberts (2011, 2020) notes, “race is not a real biological creation, but it is a real political system—invented to enforce racism, in part by claiming Black people are naturally less human”; race is a political and social system produced via social relations, not biological processes. In this section, we delineate why geneticists do not consider race a biological variable and describe the consequences of the bias in genetic research toward European ancestry populations while still cautioning that the misinterpretation of genetic research can reinforce biological narratives justifying everything from disparate medical care to white supremacy (Roberts and Rollins 2020). Throughout, we highlight the work critical of that science (Bliss 2012; Lee 2009; Panofsky and Bliss 2017; Roberts 2011), as well as sociogenomics research (Lee et al. 2018; Mills, Barban, and Tropf 2020).

### *Why Is Race Not a “Relevant Biological Variable”?*

Existing human genetics research, including sociogenomics, groups people according to ancestry rather than race. Although most geneticists and social scientists acknowledge that “there is some imprecise correlation between self-identified race and patterns of genetic variation,” if one starts from the genetic science, focusing on racial group genetic variation is illogical (Frank 2015:57; Fujimura et al. 2014). Given what is known about ancestry, it makes little sense to examine racial group genetic variation, especially in the American context. People share a social and political category, not a biological category (Morning 2014). Indeed, the

American Society of Human Genetics (2018:636), the primary professional organization representing human geneticists, notes that,

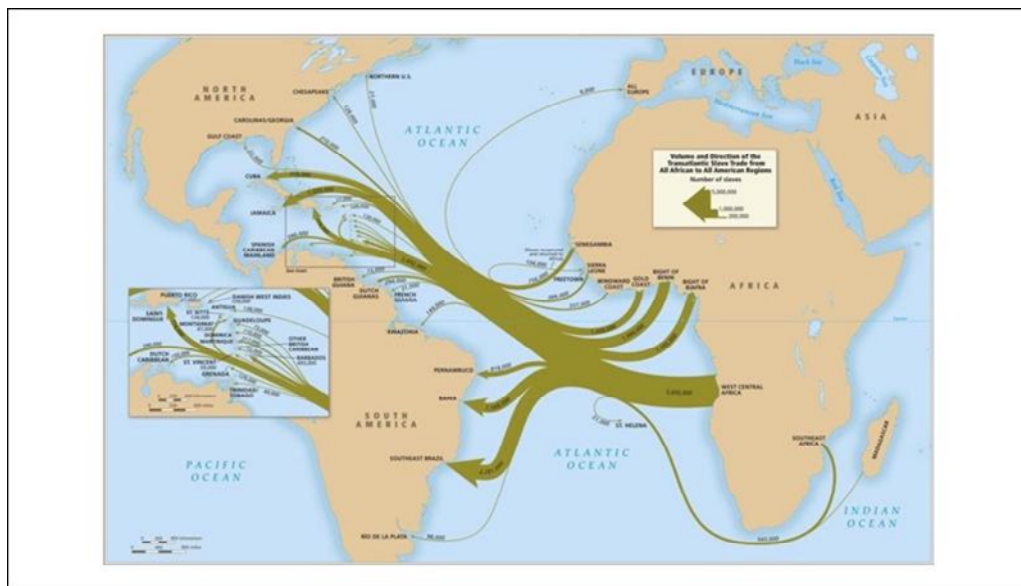
race itself is a social construct. Any attempt to use genetics to rank populations demonstrates a fundamental misunderstanding of genetics. [Ancestry] analyses are providing increasingly accurate ways of helping to define individuals’ ancestral origins and enabling new ways to explore and discuss ancestries that move us beyond blunt definitions of self-identified race.

Ancestry, rather than race, will give us tools to understand links between genetic variation and health outcomes. So what is the evidence that has led most geneticists and sociogenomics researchers to clearly state race is not a biological variable?

Genetic science points to ancestry, not race, which broadly reflects a group of people who share both a set of ancestors and a set of genetic variants (Mathieson and Scally 2020). Contemporary human genomes resulted from migration patterns, genetic reshuffling and recombination, and natural selection. Archeological and anthropological work has shown how human dispersal began approximately 2 million years ago out of Africa (Mills et al. 2020). Different migration patterns resulted in groups being dispersed to what is now known as Australia, North America, Europe, and Asia. This human dispersal of populations impacted the genetic diversity between populations, with the greatest genetic variation in sub-Saharan Africa because it houses the oldest populations (Gurdasani et al. 2015). When humans migrated to new regions, they took progressively smaller amounts of genetic variation in the gene pool with them, and each new population had less time to accumulate new mutations. As a result, European ancestry populations, for example, have more limited genetic variation.

The way geneticists examine and categorize genetic variation among population groups is by population structure patterns, which determine how populations are divided due to genetic admixture. The most common way to estimate and detect this population stratification is via principal component analysis, a well-known statistical technique that emphasizes variation and underlying patterns in the data. In populations with a relatively homogeneous background, or limited admixture, genes mirror geography (Novembre et al. 2008).

One important reason it makes little sense to group, in this case, Americans, by race is their highly varied ancestral or population background.



**Figure 1. Mapping the Slave Trade.**

Source: [slavevoyages.org](http://slavevoyages.org); Eltis, David, and David Richardson, eds. (2010). *Atlas of the Transatlantic Slave Trade*. New Haven, CT: Yale University Press.

For example, grouping Africans into a single group given the (relatively) high levels of genetic variation between populations on the African continent is largely nonsensical; the sequencing of even two people from adjacent villages of Khoi-San bushmen are as different from each other as any two European or non-African ancestry individuals (Mills et al. 2020). The same would hold for many Asian and African American populations.

From a biological viewpoint, grouping Americans by “race” based on underlying genetic factors makes little sense given the extreme ancestral diversity. As Figure 1 shows, the ancestral origins of black Americans whose ancestors were enslaved is complicated, coming from many parts of the continent (Lovejoy 2011). One must further account for immigration into the United States beyond this period, with people coming from most parts of Africa and every other part of the world from Asia to South America. Furthermore, currently, about 10% of black Americans are foreign born (Pew Research 2018). In addition, most self-identifying black Americans have substantial European ancestry, and many white-identifying Americans have non-European ancestry (Mersha and Abebe 2015). This does not even begin to account for the complexities associated with Latino-identifying Americans.

Overlaying ancestry groups onto race categories is especially illogical because sorting out ancestry differences is challenging enough. One challenge is that shared genetics among populations today, within a specific geographic location, reflect shared ancestry from another part of the world hundreds of generations ago (Skoglund and Mathieson 2018). Many ancestry-identifying tools assume that current populations accurately reflect past ancestry. The complications with this assumption would not surprise social scientists who study migration. Recent ancient DNA analyses suggest that populations have moved and mixed substantially since the ice ages (Reich 2019). Relationships between geography and genetic similarity are especially challenging at subcontinental levels, largely because of uncertain migration patterns over the past few thousand years. If anything, DNA analyses based on fossils “emphasizes discontinuity, with many populations experiencing multiple episodes of replacement over the past few thousand years” (Mathieson and Scally 2020:4). Nonetheless, new discoveries based on fossils, anthropological research, and genotyping of populations around the world are rapidly expanding knowledge (Nielsen et al. 2017).

How to measure and classify geography is also challenging. At a basic level, existing methods to identify genetic ancestry infer it indirectly by

examining genetic similarity either among those who report similar ancestry backgrounds (e.g., from Ireland) or, more commonly, to reference samples of known ancestry, which we discuss in the following. As Mathieson and Scally (2020:5) note, “recent social, economic and historical factors can distort the picture of ancestry that arises from genetic similarity.”

One common critique is the emphasis on continent as a geographic organizing principle, which can falsely blur racial and geographic difference (Panofsky and Bliss 2017). Indeed, genetics researchers have inconsistently categorized “geography.” A study of all recent genetic discoveries found 212 unique terms used to classify individuals but often mixed classifications by terms related to region, country, race-ethnicity, and ancestry (Mills and Rahal 2019). Scientists, however, are working to better standardize how ancestry is categorized in genetic research (Mills and Rahal 2020). This is especially a problem because of relatively limited genotyping and data on non-European ancestry groups. This is striking because, as already noted, of the more limited variation across European population groups (Kim et al. 2018).

All these issues, from significant ancestral diversity within geographic regions like Africa, unavailable detailed ancestral data, and difficulties measuring ancestry, make linking self-reported racial categories to genetics error ridden. A 2014 *Demography* article, which overlaid self-reported race with genetic markers, illustrates the point (Guo et al. 2014). The authors found a rough overlay between self-reported race and broad genetic continental classifications (Africa, Asia, and Europe) with high levels of admixture (for a review of similar prior work, see Mersha and Abebe 2015). They could not, however, explore considerable ancestral heterogeneity within these large geographic regions, which undercuts the utility of the analysis given the importance of precise geographic ancestry for shaping genetic variance and related health and behaviors.

Finally, biological anthropologists and geneticists make subtler points. Ancestry can be structured by the environment; sociocultural categories and practices can actually shape genetic patterns (Goodman, Moses, and Jones 2019). For example, the genetic variant for sickle cell disease, which is also protective against malaria, is more common in populations that farm yams because it requires more deforestation, which increases the amount of standing water for mosquitos to breed and thus increases malarial risk (Durham 1991).

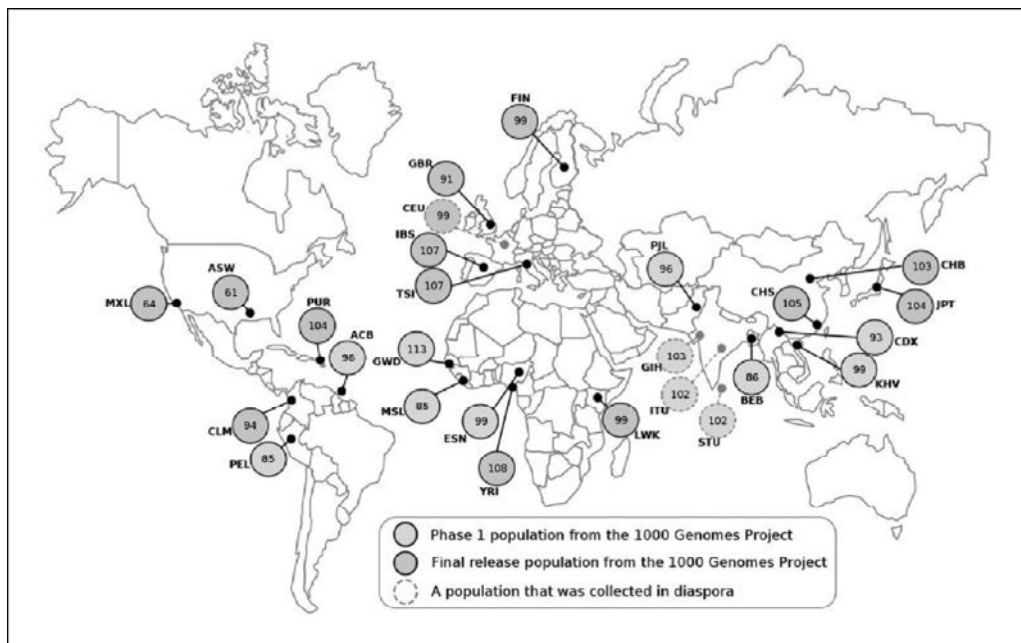
Next, we detail the exclusion of non-European ancestry groups from much of genetics research and the consequences of that exclusion.

### *The Dominance of European Ancestry Populations in Genetics Research*

Genome-wide association studies (GWASs) generate the results used by most sociogenomics and genetics researchers, who apply these results to study diseases and behavior. Simply, this is a statistical search across the genome examining each genetic variant (or region) to find a statistical relationship between a single-nucleotide polymorphism (SNP; the most common form of human genetic variation) and an outcome. Genetic variants isolated from GWASs are then used for further analyses. Nearly all complex diseases and behavioral outcomes are polygenic; they involve a wide array of genetic variants. GWASs allow scientists to identify these variants and then summarize them into a single polygenic score. Over the past 10 years, poly-genic scores have explained a growing amount of variation in disease outcomes, although socioeconomic conditions continue to explain more variation (Khera et al. 2018).

Between 2005 and 2020, 88% of GWASs were conducted on European ancestry populations. Moreover, representation is not improving over time, especially among African ancestry populations. One of the most imbalanced areas is cancer research. In 2019, European-ancestry groups made up 96% of subjects compared to 0.11% African, 0% African American and Afro-Caribbean, and 0.5% Hispanic or Latin American groups (Mills and Rahal 2020).

This bias partly results from the exclusion of non-European ancestry groups from reference panels, which are used to impute data to conduct GWASs (Mills et al. 2020). Researchers typically do not conduct whole genome sequencing for each participant. Genetic data are imputed by inferring linked or neighboring alleles of SNPs in one region due to linkage disequilibrium. Most reference panels do not include all populations (Mathieson and Scally 2020). For example, the 1000 genomes reference panel, a commonly employed reference panel, includes 2,504 individuals representing 26 populations around the world. The project acknowledges it does not reflect all populations, especially from Asia and Africa (Sudmant et al. 2015). Figure 2 provides the history and geographic coverage of these samples. Nonrandom sample selection, or convenience samples, are another weakness even



**Figure 2.** History and Coverage of the 1000 Genomes Data.

*Note:* YRI = Yoruba in Ibadan, Nigeria; JPT = Japanese in Tokyo; CHB = Chinese in Beijing; CEU = Utah residents with ancestry from northern and western Europe; LWK = Luhya in Webuye, Kenya; MKK = Maasai in Kinyawa, Kenya; TSI = Toscani in Italy; PEL = Peruvians in Lima, Peru; GIH = Gujarati Indians in Houston; CHD = Chinese in metropolitan Denver; MXL = people of Mexican ancestry in Los Angeles; ASW = people of African ancestry in the South Western United States.

*Source:* Oleksyk, Brukhin, and O'Brien (2015).

when populations are represented (Cavalli-Sforza 2005).

Adding ancestry groups better distinguishes environmental from genetic influences. To conduct GWASs on complex outcomes, we must also have very large samples of each ancestry group or at least groups with similar mixed ancestry, but most data sets are dominated by European ancestry samples. Using classic techniques when conducting a GWAS on multiple ancestry groups without stratifying by ancestry/population group risks population stratification or the conflation of environmental and genetic effects (Peterson et al. 2019). For example, one might find allele differences that correlate with drinking wine and eating cheese in a hypothetical sample that includes largely French populations but that reflects cultural rather than genetic differences (Hamer 2000).

Ancestry group exclusions also mean that one cannot identify genetic factors that may vary across populations even though nearly 85% of genetic variation is within population groups (Tropf et al. 2017). GWAS only detects variants that are

common in the study population. If one only includes European-ancestry groups, variants that vary across ancestry groups will not be detected. Furthermore, applying findings on genetic disease risk in one population to another population can misestimate disease risks because the tool is based on European sequencing samples (Kim et al. 2018). Because differences in allele frequencies vary across ancestry groups, this results in ascertainment bias, or selection, because in some cases, alleles discovered in GWASs may have a systematically higher or lower frequency in European-ancestry populations (Lachance and Tishkoff 2013). SNP ascertainment bias also occurs because genotyping arrays contain a biased set of preascertained SNPs (Lachance and Tishkoff 2013). Moreover, there is evidence that polygenic scores developed in ancestry-specific populations have limited portability across populations (Duncan et al. 2019).

Largely European-ancestry data also narrow the ability to understand gene–environment interplay. It is difficult to study gene-by-environment interactions among individuals in Western and wealthy

nations who have a non-European ancestry background. Without sufficient ancestral information, including these individuals in analyses introduces population stratification (Peterson et al. 2019). Furthermore, limiting analyses to Western and wealthy nations profoundly limits the ability to understand gene–environment interplay by limiting environmental variation.

Researchers are trying to address this with new methods allowing for more ancestrally diverse groups in GWASs as well as broader genotyping across populations (Peterson et al. 2019). But change is slow and does not match the pace of genetic research and increasingly diverse societies.

### *Scientific Racism and Its Current-Day Consequences*

Given that the human genome was mapped just 20 years ago, some of these challenges, like measuring ancestry, are unsurprising. No scientific enterprise, especially those at the cutting edge, is devoid of error. Indeed, discussion and debate are the only ways to make scientific progress. The problem, however, is that genetics is laden with a legacy of scientific racism, and there is evidence of its resurgence (Martschenko, Trejo, and Domingue 2019). Sociologists have argued that imprecise discussions of ancestry and an overemphasis on population differences can reinforce views that race is biological (Duster 2004; Fujimura and Rajagopalan 2011; Fullwiley 2008; Morning 2011; Panofsky and Bliss 2017; Phelan, Link, and Feldman 2013; Roberts 2011; Wailoo, Nelson, and Lee 2012).

A recent *New York Times* editorial by David Reich (2018a), a world-renowned specialist in ancient genomics, demonstrates how genetic science can be (mis)interpreted in the broader public. He noted there may be meaningful differences across ancestral populations in the genetic risk for disease or behavioral outcomes. He then indicated that because ancestral populations can overlap with race—which he acknowledged is a social construct—there may be genetic race differences. He goes on to say, “I have deep sympathy for the concern that genetic discoveries could be misused to justify racism. But as a geneticist I also know that it’s simply no longer possible to ignore average genetic differences among ‘races.’” The editorial raised an uproar, leading to a subsequent interview in which Reich tried to clarify his views (Kahn et al. 2018; Reich 2018b).

Why the uproar? Despite stating that race and ancestry were distinct concepts, Reich (2018a)

concluded by then noting genetic “race” differences. Notably, in his recent book, he used the term *populations*, not *race*, in the same sentence (Holmes 2018; Reich 2019). But an ironic use of quotes around *race* to mean population would not be clear to your average *Times* reader. Sociologists who argued that ancestry, in practice, could be just another term for race saw this editorial as validating that fear (Kahn et al. 2018).

Also concerning was Reich’s supposition that “while race may be a social construct, differences in genetic ancestry that happen to correlate to many of today’s racial constructs are real,” particularly linking to behavioral and health outcomes (Kahn et al. 2018; Reich 2018a). As previously detailed, it is a highly speculative scientific claim and one on which it makes little sense, scientifically, to expend one’s energy. Indeed, one of the two studies he cited has already been disproven. He noted that height differences between southern and northern Europeans are explained by genetic differences. The first assumption he makes is that southern and northern Europeans constitute different races, pointing to the highly political and social ways in which we define race. Second, there is evidence that these results were confounded by unmeasured environment (Sohail et al. 2019).

Speculating a connection between genetics, race, and complex outcomes in the existing political and social order, particularly in the United States, has profoundly different implications than, for example, speculating a relationship between genetics, excessive carrot consumption, and heart disease. Indeed, incorrect assumptions of biological race difference have contributed to racial disparities in medical care. A recent *New England Journal of Medicine* article noted that “despite mounting evidence that race is not a reliable proxy for genetic difference, the belief that it is has become embedded, sometimes insidiously, within medical practice” (Vyas, Eisenstein, and Jones 2020:874). For example, the American Heart Association has different cut points for cardiac risk scores by race; one consequence, as shown in a study of Boston emergency rooms, is that black and Latinx individuals were less likely to be admitted to cardiology services when they had heart failure (Eberly et al. 2020). These race adjustments affect racial disparities in care widely, including in nephrology, obstetrics, and urology, in addition to critical procedures like organ transplants (Vyas et al. 2020).

The public also does not understand the distinction between race and ancestry. According to a 2018 poll, around 34% of Americans believe

genetics fully determines your racial identity (Tillery 2018). One recent study even found that more than 20% of Americans attribute differences in socioeconomic conditions between black and white Americans to genetics, and that this view reinforces racial prejudice (Byrd and Ray 2015; Morning, Brückner, and Nelson 2019). Indeed, white people are more likely to believe genetic explanations for racial inequalities, which reduces their support for policies that could address racial inequality resulting from systemic racism (Byrd and Ray 2015). Furthermore, white supremacists have actively misused genetic research, especially linked to ancestry, to make claims about white superiority (Panofsky and Donovan 2019). They have even infiltrated universities and utilized seemingly legitimate open-access journals to peddle racial genetic essentialism (Saini 2019).

## GENETICS, SOCIAL ENVIRONMENTS, AND EDUCATIONAL ATTAINMENT

Like race, education has been central to sociology of health. It is a fundamental cause, shaping morbidity and mortality across time and over the life course (Herd 2010; Herd, Goesling, and House 2007; House et al. 1994; House, Lantz, and Herd 2005; Link and Phelan 1995). But how can this perspective be reconciled with research showing genetic influence on educational attainment?

We discuss gene–environment interplay in determining educational attainment; the role for sociologists given our theoretical, substantive, and methodological expertise; and the implications of this science given the history of eugenics. Sociologists have been rightfully wary of reductionist genetic explanations for educational disparities. We argue that engaging in this research will help clarify how the social and biological interact to affect inequality. Perhaps more importantly, however, sociologists engaging will help ensure genetics research is not reductionist.

### *Understanding Genetics, Environment, and Educational Attainment: Twin Research*

Social scientists do assume that genetics, in part, shapes educational attainment, even when grappling with eugenics ideology. The popularity of sibling models (e.g., Warren, Sheridan, and Hauser 2002) and exogenous shocks, like mandatory schooling

laws (e.g., Glymour et al. 2008), reflect an assumption that genetic endowments shape, in part, educational outcomes. But while showing causal effects of environment on educational attainment, these methods cannot explore gene–environment interplay.

Behavioral genetics research, rooted in psychology, has shown genetic influence but also evidence of environmental modification. Empirically, researchers have used twin models to estimate heritability: the fraction of the variance in an outcome, like educational attainment, that is due to genetic variation between individuals in that population. Until recently, twin studies dominated genetic studies of educational attainment (see Branigan, McCallum, and Freese 2013). On average, about 40% of the variance in educational attainment was accounted for by genetics. Importantly, estimates ranged from 18% to 76%, implying heritability varies across environments or populations (Branigan et al. 2013).

But similar to sibling and exogenous shock models, twin studies struggle to examine gene–environment interplay. They typically have small and select samples, weakening their generalizability and statistical power. Twins also differ from singletons on language development, personality, and internalizing behavioral problems, a problem for education analyses (Christensen and McGue 2020). Furthermore, environmental measures in twin studies better capture family environments than broader social and institutional environments (Boardman, Daw, and Freese 2013).

### *Understanding Genetics, Environment, and Educational Attainment: Polygenic Scores*

Polygenic scores offer an alternative approach to examine gene–environment interplay. They are generated from GWASs, discussed previously, which analyze relationships between the most common form of genetic variation (SNPs) and behavioral and health outcomes. They are polygenic because single variants rarely have a large influence on a single outcome. Polygenic scores are a summary score for each individual; it is a weighted average of the strength of the association between each genetic variant and the outcome. Robust polygenic scores have been developed for outcomes ranging from height and BMI to psychiatric disorders and fertility (Duncan et al. 2019).

A polygenic score for educational attainment was targeted quickly for a few reasons. First, education is

the most studied social input and outcome in the social sciences. Second, it could advance research on gene–environment interplay with its inclusion in representative, randomly selected, longitudinal studies with comprehensive measures of environment. Finally, it is common in most data sets, allowing for large sample sizes to detect genetic variant influences. Three educational attainment polygenic scores were developed between 2013 and 2018, each composed of more variants as the sample sizes grew (Lee et al. 2018). Like most GWASs, it was conducted on European-ancestry individuals, including 71 different studies ranging from 23andMe to the Health and Retirement Study. It now explains about 11% of the total variance, which nears classic predictors such as one parent’s educational attainment (Lee et al. 2018).

New evidence has shown that social conditions modify genetic influences on educational attainment. For example, gender modifies the influence of the polygenic score on educational attainment, with variance over the life course and across cohorts, supporting findings from twin meta-analyses (Branigan et al. 2013; Herd et al. 2019; Silventoinen et al. 2020). Furthermore, the score’s influence on academic achievement varies depending on school resources. Students with low poly-genic scores from advantaged schools were less likely to drop out of math than were similar students from less advantaged schools (Harden et al. 2020). We detail more of this work in the next section.

Yet there are methodological challenges that may infer the score is overestimating genetic influence (Akimova et al. 2020). First, the education polygenic score is not a direct effect. Like other polygenic scores, it is a black box. Thousands of genetic variants make up the score, many correlated with other behavioral and environmental factors. For example, the polygenic score for educational attainment predicts schizophrenia, personality, and cognitive functioning (Bansal et al. 2018). Relatedly, there is overlap between the education polygenic score and other outcomes, like age at first birth, which has an approximately 70% to 74% overlap with genetic loci found with educational attainment (Mills and Tropf 2020). This is the ubiquity of pleiotropy for complex outcomes, which is where one genetic variant influences more than one outcome (Wedow et al. 2018). This overlap implies the scores may include environmentally mediated pathways.

Recent studies have also shown that polygenic scores may overstate genetic influence for other

reasons. Those generated on samples of unrelated individuals are susceptible to population stratification bias; polygenic scores may capture environment, such as social norms or educational policies, leading to overestimates of heritability (Mills and Tropf 2020). Although most scores are affected, the education polygenic score appears to be more so (Brumpton et al. 2020). The score also likely picks up environment due to genetic nurture—how one’s parents’ genetics shapes the environment in ways that influence outcomes (Selzam et al. 2019). But these effects can be managed with the use of family and sibling models, for example (Selzam et al. 2019).

Alternately, the education polygenic score may underestimate genetic influence. There also may be “missing heritability.” The education polygenic score explains approximately 25% to 30% of the estimated heritability suggested by twin studies, although many believe twin studies overestimate heritability (Kim et al. 2018; Nolte, Troph, and Snieder 2019). Others argue that GWAS techniques hide heritability due to population differences, like living in different countries (Tropf et al. 2017). Despite proposed new estimation approaches, like rare variant inclusion in GWASs, there are not firm answers.

### *Understanding Genetics, Environment, and Educational Attainment: The Role for Sociology*

Existing research examining how gene–environment interplay shapes educational attainment is promising, but challenges remain. Here, we outline sociology’s theoretical and methodological tools, which can enhance genetics research with closer attention to environment, while also deepening sociology’s understanding of the relationship between the social and biological.

Sociologists view environment as social and structural, and so psychology’s dominance and sociology’s relative absence in the genetics of intelligence and education research led to a more individualized view of environment (Freese 2008; Herd et al. 2019). For example, some psychologists view IQ and educational attainment as interchangeable (Plomin 2018). Sociologists, of course, know “the difference between a psychological trait and a life course attainment, as the latter occurs only as the product of an extended chain of interactions between individual behavior and environmental response” (Herd et al. 2019:43). Indeed, it was sociologists that showed, in a meta-analysis of twin studies, the much larger role of environment in the



heritability of education versus IQ (Branigan et al. 2013).

Sociologists emphasize that genetic influence must be understood in context, such as history and cohort, the life course, and social forces like gender; the application of the sociological imagination can clarify how social and historical conditions modify genetic influences on educational attainment (Boardman et al. 2013; Freese 2008; Herd et al. 2019). The influence of genetic factors on educational attainment is filtered, altered, and shaped by broader complex environments. This contrasts with psychology's sometimes more "atomistic" view of environment, which is family focused, partly reflecting the dominance of twin studies (Boardman et al. 2013; Plomin 2018).

Recent research, compared to decades of twin studies, has vastly expanded empirical evidence that actually shows how social context modifies genetic influence. Studies have examined how institutions, such as schools, school funding, geography, and educational policies, as well as broader social forces, such as cohort, the life course (or time), and gender, modify the influence of the education polygenic score on educational attainment (Conley et al. 2016; Domingue et al. 2018; Harden et al. 2020; Herd et al. 2019; Liu 2018; Schmitz and Conley 2017; Trejo et al. 2018; Ujma et al. 2020; Wedow et al. 2018). Researchers have even tested and debunked long-standing theories that growing stratification in the United States is largely a function of genetic inheritance (Conley and Domingue 2016).

Greater use of the education polygenic score and gene-by-environment analyses, more broadly, may also improve the understanding of social impacts. In short, social influences may only become evident once one integrates genetic data (Johnson, Sotoudeh, and Conley 2020). For example, a direct social effect may be washed out because its dominant influence is only in a small genetically high-risk group. In the following section, we detail new methods that can help detect these kinds of effects.

Going forward, however, we must address the polygenic score's methodological challenges to advance gene-environment interplay research. Sociologists can and have contributed to these innovations. A key challenge is the risk for confounding due to endogeneity between the polygenic score, the environment, and the outcome. There are different ways to address this. For example, exogenous policy changes can be useful tools for both navigating endogeneity while also contributing

relevant evidence to inform structural change via policy choices. Indeed, a recent study employed polygenic scores and mandatory schooling policies in the UK, demonstrating a causal impact of schooling on health and mortality (Davies et al. 2018). The use of family-based models, a long-standing practice in sociology, also reduces this error.

Another concern is that the current methods, while allowing tests on how environments modify an individual's genetic propensity "to attain that same outcome," are not adept at testing plasticity or "an individual's propensity towards variability in an outcome" (Johnson et al. 2020:1). Individuals may have an underlying genetic propensity for being highly responsive to environments. Recent research is tackling this (Johnson et al. 2020).

Sociologists are also well poised to unpack the black box of a polygenic score like educational attainment, clarifying its downstream environmental mediators, given our mixture of rich data and already well-understood environmental influences on educational attainment. Akimova and colleagues (2020) pointed to key methodological issues and possible solutions to unpack these mediating pathways.

We also need the right data to test how environments modify genetic influence. Nonrepresentative and nonrandomly selected study populations, common in medical research, are a central challenge. For example, the UK Biobank is commonly used for genetic and gene-by-environment analysis. It is, however, highly selective on health and higher socioeconomic status, even including significant geographic clustering (introducing population stratification), reflecting clinic-based survey assessment and lack of nonresponse conversion (Fry et al. 2017). Interestingly, its large sample size is undermined by its selection, limiting environmental variation and reducing statistical power (Domingue et al. 2020). Narrow social and economic measures are also problematic. Social scientists, however, have been adding genetic data to large longitudinal studies with randomly selected and representative analytic samples, including rich and high-quality environmental measures (Herd et al. 2014; McQueen et al. 2015; Sonnega et al. 2014).

### *Examining How Education Moderates Genetic Risk for Health and Reproductive Outcomes*

One also can learn how social conditions modify the relationship between polygenic risk for disease and disease outcomes. Numerous robust polygenic

scores for reproductive outcomes, like age at first birth and number of children, and disease outcomes, including BMI, coronary artery disease, atrial fibrillation, type 2 diabetes, inflammatory bowel disease, and breast cancer, have been developed (Khera et al. 2018). Indeed, cohort and educational attainment moderate the association between the polygenic score for BMI and BMI outcomes (Conley et al. 2016; Liu and Guo 2015). UK schooling reforms led to 14% of students completing an additional year of school, which offset the genetic risk for obesity, body size, high blood pressure, and lung problems (Barcellos, Carvalho, and Turley 2018). Furthermore, higher childhood socioeconomic status offset genetic risk for early reproductive onset and higher externalizing behavior (Mills et al. 2020).

### *Social and Political Critiques: Concerns about Genetic Determinism*

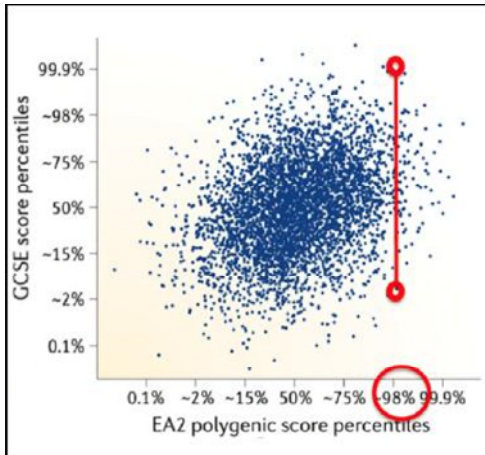
Yet even with possibility, risk remains. Will this research reinvigorate eugenics and genetic determinism, the view that behaviors are directly controlled by genes with minimal influence of environment (Martschenko et al. 2019; Roberts 2015)? Early twentieth-century eugenics thinking was widely embraced by scientists and the public. It influenced policymaking as a justification for nativist immigration policies, policies that limited women's labor force participation, and racist policies, including the forced sterilization of black women and disabled people (Leonard 2017). Eugenics lost influence following World War II and the atrocities in Nazi Germany. Geneticists, social scientists, and the broader public rejected the logic that genetics was completely deterministic or that a better society should be "bred" (Saini 2019). But while losing its luster, support remained. Indeed, forced sterilization continued in some states until the 1970s (Hansen and King 2013). Even in academia, a small subgroup of scholars retained the basic ideologies of early twentieth-century eugenics.

Concerningly, there is evidence of a resurgence in this thinking (Saini 2019). Recent books by Charles Murray (2020) and Robert Plomin (2018) raise concerns regarding how genetic research, especially examining genetic influence on educational attainment, could reinvigorate genetic determinism. They serve as case studies to illustrate how genetics research has been extrapolated in ways not supported by evidence to make bold claims about genetic determinism that filter out into the public and policy discourse.

Charles Murray's (2020) book, *Human Diversity: Biology of Gender, Race, and Class*, tries to link race and the genetics of cognition. The leaps of logic that Murray used to conclude that there are race differences in IQ driven by genetic differences are, scientifically, easily refuted. Yet despite behavioral geneticists thoroughly debunking the claims, Murray received significant attention in the media (Turkheimer, Harden, and Nisbett 2017). Moreover, the media coverage often got the basic science wrong. Some journalists outright supported Murray's claims, completely ignoring the expertise of the most influential behavioral geneticists studying cognition (Turkheimer et al. 2017). The deference to Murray was especially striking given that Murray is a political scientist with a dubious history in the use of statistics and no genetics expertise. Even those who saw the error of Murray's ways still misunderstand basic scientific facts. For example, William Saletan (2018), a writer for *Slate*, accurately critiqued the connection between research on the "genetics of intelligence" and the "genetics of race" but did not understand that there is no genetics of race.

Robert Plomin's (2018) book, *Blueprint*, points to a different dilemma regarding the use of genetics in debates about educational attainment and intelligence and especially how they are used in policy discussions. Unlike Murray, Plomin is an expert in behavioral genetics. His empirical work, particularly based on twin studies, is well respected and highly cited. *Blueprint*, however, which is intended for popular audiences, makes some very bold claims. Plomin (2018:vii) states that genetics will allow individuals to not only "understand who we are" in the present but also predict "who we will become" in the future. He adds that DNA can "tell your fortune from the moment of your birth, it's completely reliable and unbiased" (p. vii). If one notes the sleight of hand, his claim is individual. He is not saying that genetics explain significant amounts of variation at the population level; he is claiming that people can, or will be able to, look at an individual's DNA and predict who that person will become. People cannot do that—and all available evidence indicates that people will never do that in a meaningful way. What is striking is that his own work—cited throughout *Blueprint*—does not actually support this kind of genetic determinism.

To give a specific example of this overstatement, Plomin claims that the educational attainment polygenic score could be used for "precision education" (Plomin and von Stumm 2018), but educational polygenic scores cannot predict individual-level performance (Morris, Davies, and Smith



**Figure 3.** Toward Precision Education?

Note: Adapted from Plomin and von Stumm (2018).

2020). Figure 3 uses Plomin and von Stumm's (2018) correlation of the educational attainment score by middle school score percentiles; if someone has a polygenic score at the 98th percentile, they could score anywhere between the 2nd and 98th percentiles, which is hardly precision (Mills et al. 2020).

Science, rather than being translated rationally, is often used to make political claims (Roberts 2015; Roberts and Rollins 2020). In public forums and the popular press, Plomin frequently overstates or muddies existing evidence. For example, in a recent forum intended to inform UK policymaking, he claimed that schools do not “make a difference in [children's] educational achievement” and that education polygenic scores could eventually be used for “personalized” education (Plomin 2019). These are demonstrably false and overstated claims (Jackson 2018; Mills et al. 2020). Yet genetic determinism, as applied to education, has infiltrated policy debates in the UK. In 2013, a key advisor to the English secretary of education, after extensively citing twin research by psychologists like Plomin, claimed that schools or teacher quality had little influence on educational outcomes (Merrick 2013).

One issue, easily misconstrued by policymakers, is genetic nurture; environmental effects may have genetic origins (Selzam et al. 2019). Indeed, Plomin's claims about the dominance of genetic influence on educational attainment rests on genetic nurture. For example, if a parent reads to a child, the child will be more likely to read, but part of that environmental influence is genetic—the parent reads more due to his or her own genetics. But by

stating the effect is really genetic, one risks obfuscating the counterfactual; if the parent stops reading to the child, even one with genetic potential, he or she will still be less likely to read. This distinction is critical from a policymaking perspective.

What we have detailed is how research on the genetics of educational attainment could intensify social inequality (Roberts 2015). We believe that failing to engage increases these risks. Rather than ignoring genetic influence, which contradicts most scientific evidence, sociologists of health instead may demonstrate how genetic influence is unextractable from environment.

## CONCLUSION

Sociogenomics research has significant relevance for sociology, including both risks and rewards. We now point to possible actions to reduce the risk, including how sociogenomics research is used and understood in the broader public domain, changes to the research itself, and collective institutional responses.

### *We Need to Be Cautious about How We Communicate Our Research*

Sociogenomics researchers must be aware that research findings are not interpreted in a vacuum, but in a social and political context. For example, the distinction between race and population or population geneticists' consensus that race *is not* a biological construct is not common knowledge (Tillery 2018). This should shape how scientists communicate research. David Reich loosely conflating race and population is a powerful example of not considering the public's scientific knowledge. In contrast, while discussing her study on lacking diversity in genetic studies (Peterson et al. 2019), Alice Popejoy explicitly noted that people “shouldn't get the impression that health disparities are driven by differences in genetic structure between ethnic groups. Environment matters and widespread systemic and structural racism that exacerbates environmental effects are more important” (Lambert 2019).

Scientists should be cautious about making claims that go beyond existing evidence (Martschenko et al. 2019). For example, David Reich made a highly speculative claim that population differences that map onto socially constructed race categories will likely be found. Because most Americans already believe that race has a strong genetic basis, his approach misleads rather than informs (Kahn et al. 2018). Furthermore, given the myriad of ways in which geneticists have gotten the science

wrong—and the racist consequences of that fact—scientists must exhibit far more humility regarding speculative claims. Indeed, experiments demonstrate that beliefs in racial genetic difference increase racial prejudice (Morning et al. 2019). Scientists making speculative claims about racial genetic differences cause harm.

### *We Need to Improve Public Understanding of Genetic Research*

Improving the public's scientific knowledge is another way to address risk. A recent study provides a promising approach. When people overestimate genetic differences between groups, they exhibit more racial bias (Morning et al. 2019). What happens when people are provided with accurate information about genetic diversity? A replicated study showed that teaching adolescents a week's worth of material on human genetic variation significantly decreased bias (Donovan et al. 2019).

### *How the Science Needs to Change*

Limited representation of non-European ancestry populations in genetic studies also reflects, in part, racist ideologies. The practice is so common across scientific research that there is a term for it: *WEIRD* (Western, educated, industrialized, rich, and democratic) samples, which also applies to *WEIRD* scientists. Many researchers and professional organizations have called for more inclusive samples and more inclusive research communities (Bentley, Callier, and Rotimi 2017; Kim et al. 2018; Mills and Tropf 2020). For example, H3Africa is a joint initiative across government and nonprofit groups to increase genetic research across the African continent, including data collection. Yet the pace of new *WEIRD* research is outpacing efforts like this one.

The inclusion of communities affected by research must be included in the research process to ensure equitable outcomes. For example, community-based participatory research models have successfully engaged indigenous communities within the United States (Claw et al. 2018). These communities want a say in how their genetic data are used and translated. Big data studies circumventing these approaches fail at study recruitment (Tsosie, Yracheta, and Dickenson 2019).

Scientific communities must be more representative. Genetics research, including sociogenomics research, remains dominated by white men (Bliss 2018). A recent analysis of GWASs found that women were senior author on just one quarter of

those publications (Mills and Rahal 2019). White Americans and Europeans also dominate this research, with very few black, indigenous, and Latinx scientists, which affects the science. Indeed, minority population geneticists have been disproportionately working to diversify genetic data samples (Guglielmi 2019). One model is the Summer Internship for Indigenous Peoples in Genomics program run by molecular anthropologist Ripan Malhi, which trains people from indigenous communities on genetics.

The inclusion of groups being studied in the actual research process is critical. To highlight another example, a study into the genetic architecture of same-sex populations (Ganna et al. 2019) was explicit to include gay men and was careful to clearly detail in publicly accessible ways the meaning of this research, as well as the ways it could be misused (Phelps and Wedow 2019).

Finally, as we detailed, sociologists have a unique skill set to examine how environments modify genetic influence. But training on how to do gene-environment research is limited. Although there are sociogenomics training workshops, trainings focused specifically on gene-by-environment research would be particularly valuable.

### *Institutional and Collective Responses by Social Scientists*

The aforementioned recommendations are not comprehensive, nor do they necessarily represent the consensus of the broader social scientific community. Consequently, given sociogenomic's rapid growth and the related ethical and scientific risks, a broader collective and institutional response is needed. Indeed, geneticists, via their professional organizations, have developed everything from shared belief statements—such as race is a social, not a biological, construct—to ethical guidelines around the use of devices like CRISPR. The social science community should also engage in these collective and institutional responses. The convening of a National Academy of Sciences Panel would be a constructive way to begin developing collective practices and standards. Although there are numerous issues with which social scientists must grapple, we focus here on two in particular that would be excellent places to start building consensus: data use and policy implications of sociogenomics research.

Social science research has always been on the forefront of data sharing. The field's data are a public good that is shared widely and easily. Longitudinal

studies, including the Health and Retirement Study, the Wisconsin Longitudinal Study, and the National Longitudinal Study of Adolescent to Adult Health (Add Health), all now include genetic data and polygenic scores. This approach generally improves science—open access improves reliability, democratizes access, and increases data usage.

The open sharing of genetic data has risks, however. White supremacists using these data, with limited training combined with an ideological agenda, is toxic. Indeed, a recent study, published in an American Psychological Association journal, erroneously used genetic data to claim Jewish people had a genetic advantage that conferred higher levels of educational attainment. Although easily debunked (Freese et al. 2019), the journal refused a request for a study retraction (Mills and Tropf 2020). Do data providers need to think differently about how these data are shared, particularly given the lack of responsiveness, in this case, from a journal sponsored by a major professional organization?

Social scientists also must address that study findings may appear in debates within policy domains, especially education. Some individual researchers have been thoughtful about developing statements linked to research findings regarding policy implications. For example, the lead authors of the study that generated the polygenic score for educational attainment developed a FAQ for how to interpret the score (Lee et al. 2018). They collaborated with an ethicist on key points, including a clear statement noting that one could not draw *any* policy or practice conclusions and to do otherwise would be premature. The science did not support those inferences (Lee et al. 2018). Our field should develop shared norms and practices that address how sociogenomic research findings should be—or should not be—used in policy domains. Although we, as researchers, cannot always control how evidence is used, we can make clear how we think it should be used.

There are numerous other challenges, like issues with expertise in peer review, but the broader point is that the social science community should follow the lead of the genetics community and develop some collective standards and strategies for sociogenomics research as well as the dissemination of this research.

There is tremendous promise but also risk in the practice of genetic research by social scientists. Work by critical race theory scholars like Alondra Nelson, Ann Morning, Dorothy Roberts, Troy Duster, Ruha Benjamin, Duana Fulliley, Catherine Lee, Joan Fujimura, and Catherine Bliss

have profoundly influenced basic genetic research by drawing attention to how that work could reinforce false biological notions of race. Sociologists who engage in sociogenomics research are making, and can continue to make, substantive and critical contributions. Sociology can bring a unique lens to this work, with expertise in theory, methods, and substance, all of which can ultimately help to better understand the interplay between the social and biological and reduce the risk for reductionist understandings of genetic influence.

## FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding provided by ERC-Adv-2018-835079 and the Leverhulme Trust, Leverhulme Centre for Demographic Science.

## ORCID ID

Pamela Herd <https://orcid.org/0000-0002-0137-5846>

## REFERENCES

- Akimova, Evelina, Richard Breen, David Brazel, and Melinda Mills. 2020. "Heritable Environments." *bioRxiv*. doi:10.1101/2020.11.02.364539.
- American Society of Human Genetics. 2018. "ASGH Denounces Attempts to Link Genetics and Racial Supremacy." *The American Journal of Human Genetics* 103(5):636.
- Bansal, Vikas, Marina Mitjans, Casper Burik, Richard Karlsson Linner, Aysu Okbay, Cornelius Rietveld, Martin Begemann, et al. 2018. "Genome-Wide Association Study Results for Educational Attainment Aid in Identifying Genetic Heterogeneity of Schizophrenia." *Nature Communications* 9(1):3078. doi:10.1038/s41467-018-05510-z.
- Barcellos, Silvia, Leandro Carvalho, and Patrick Turley. 2018. "Education Can Reduce Health Differences Related to Genetic Risk of Obesity." *Proceedings of the National Academy of Sciences* 115(42):E9765–72.
- Bentley, Amy, Shawneequa Callier, and Charles Rotimi. 2017. "Diversity and Inclusion in Genomic Research." *Journal of Community Genetics* 8(4): 255–66.
- Bliss, Catherine. 2012. *Race Decoded*. Stanford, CA: Stanford University Press.
- Bliss, Catherine. 2018. *Social by Nature*. Stanford, CA: Stanford University Press.
- Boardman, Jason, Jonathan Daw, and Jeremy Freese. 2013. "Defining the Environment in Gene–Environment Research." *American Journal of Public Health* 103(S1):S64–72.
- Branigan, Amelia, Kenneth McCallum, and Jeremy Freese. 2013. "Variation in the Heritability of

- Educational Attainment." *Social Forces* 92(1): 109–40.
- Brumpton, Ben, Eleanor Sanderson, Karl Heilbron, Fernando Pires Hartwig, Sean Harrison, Gunnhild Åberge Vie, Yoonsu Cho, et al. 2020. "Avoiding Dynastic, Assortative Mating, and Population Stratification Biases in Mendelian Randomization through Within-Family Analyses." *Nature Communications* 11(1):3519. doi:10.1038/s41467-020-17117-4.
- Byrd, Carson, and Victor Ray. 2015. "Ultimate Attribution in the Genetic Era." *The Annals of the American Academy of Political and Social Science* 661(1):212–35.
- Cavalli-Sforza, Luca. 2005. "The Human Genome Diversity Project." *Nature Reviews Genetics* 6(4):333–40.
- Christensen, Kaare, and Matt McGue. 2020. "The Twin Representativeness Assumption." *Neuroscience and Biobehavioral Reviews* 112:374–75.
- Claw, Katrina, Matthew Anderson, Rene Begay, Krystal Tsosie, Keolu Fox, and Nanibaa A. Garrison. 2018. "A Framework for Enhancing Ethical Genomic Research with Indigenous Communities." *Nature Communications* 9(1):2957. doi:10.1038/s41467-018-05188-3.
- Conley, Dalton, and Benjamin Domingue. 2016. "The Bell Curve Revisited." *Sociological Science* 3: 520–39.
- Conley, Dalton, Thomas Laidley, Jason Boardman, and Benjamin Domingue. 2016. "Changing Polygenic Penetrance on Phenotypes in the 20th Century among Adults in the US Population." *Scientific Reports* 6:30348. doi:10.1038/srep30348.
- Davies, Neil, Matt Dickson, George Davey Smith, Frank Windmeijer, and Gerard van den Berg. 2018. "The Effect of Education on Adult Mortality, Health, and Income." *bioRxiv*. doi:10.1101/250068.
- Domingue, Benjamin, David Rehkopf, Dalton Conley, and Jason Boardman. 2018. "Geographic Clustering of Polygenic Scores at Different Stages of the Life Course." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 4(4):137–49.
- Domingue, Benjamin, Sam Trejo, Emma Armstrong-Carter, and Elliot Tucker-Drob. 2020. "Interactions between Polygenic Scores and Environments: Methodological and Conceptual Challenges." *Sociological Science* 7:465–86.
- Donovan, Brian, Rob Semmens, Phillip Keck, Elizabeth Brimhall, K. C. Busch, Monica Weindling, Alex Duncan, et al. 2019. "Toward a More Humane Genetics Education." *Science Education* 103(3):529–60.
- Duncan, Laramie, H. Shen, B. Gelaye, J. Meijsen, K. Ressler, M. Feldman, R. Peterson, and Ben Domingue. 2019. "Analysis of Polygenic Risk Score Usage and Performance in Diverse Human Populations." *Nature Communications* 10(1):3328. doi:10.1038/s41467-019-11112-0.
- Durham, William. 1991. *Coevolution*. Stanford, CA: Stanford University Press.
- Duster, Troy. 2004. *Backdoor to Eugenics*. New York, NY: Routledge.
- Eberly, Lauren, Sharlene Day, Euan Ashley, Daniel Jacoby, John Jefferies, Steven Colan, Joseph Rossano, et al. 2020. "Association of Race with Disease Expression and Clinical Outcomes among Patients with Hypertrophic Cardiomyopathy." *JAMA Cardiology* 5(1):83–91.
- Frank, Reanne. 2015. "Back to the Future? The Emergence of a Geneticized Conceptualization of Race in Sociology." *The ANNALS of the American Academy of Political and Social Science* 661(1): 51–64.
- Freese, Jeremy. 2008. "Genetics and the Social Science Explanation of Individual Outcomes." *American Journal of Sociology* 114(S1):S1–35.
- Freese, Jeremy, Ben Domingue, Sam Trejo, Kamil Sicinski, and Pamela Herd. 2019. "Problems with a Causal Interpretation of Polygenic Score Differences between Jewish and non-Jewish Respondents in the WLS." *SocArXiv*. doi:10.31235/osf.io/eh9tq.
- Fry, Anna, Thomas Littlejohns, Cathie Sudlow, Nicola Doherty, Ligia Adamska, Tim Sprosen, Rory Collins, and Naomi Allen. 2017. "Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population." *American Journal of Epidemiology* 186(9):1026–34.
- Fujimura, Joan, Deborah Bolnick, Ramya Rajagopalan, Jay Kaufman, Richard Lewontin, Troy Duster, Pilar Ossorio, and Jonathan Marks. 2014. "Clines without Classes." *Sociological Theory* 32(3):208–27.
- Fujimura, Joan, and Ramya Rajagopalan. 2011. "Different Differences: The Use of 'Genetic Ancestry' Versus Race in Biomedical Human Genetic Research." *Social Studies of Science* 41(1):5–30.
- Fullwiley, Duana. 2008. "The Biological Construction of Race." *Social Studies of Science* 38(5):695–735.
- Galton, Francis. 1869. *Hereditary Genius*. New York, NY: Macmillan.
- Ganna, Andrea, Karin Verweij, Michel Nivard, Robert Maier, Robbee Wedow, Alexander S. Busch, Abdel Abdellaoui, et al. 2019. "Large-Scale GWAS Reveals Insights into the Genetic Architecture of Same-Sex Sexual Behavior." *Science* 365(6456):eaat7693. <https://doi.org/10.1126/science.aat7693>.
- Glymour, Maria, Ichiro Kawachi, Christopher Jencks, and Lisa Berkman. 2008. "Does Childhood Schooling Affect Old Age Memory or Mental Status?" *Journal of Epidemiology & Community Health* 62(6):532–37.
- Goodman, Alan, Yolanda Moses, and Joseph Jones. 2019. *Race: Are We So Different?* New York, NY: John Wiley and Sons.
- Guglielmi, Giorgia. 2019. "Facing Up to Injustice in Genome Science." *Nature* 568(7752):290–94.
- Guo, Guang, Yilan Fu, Hedwig Lee, Tianji Cai, Kathleen Mullan Harris, and Yi Li. 2014. "Genetic



- Bio-ancestry and Social Construction of Racial Classification in Social Surveys in the Contemporary United States." *Demography* 51(1):141–72.
- Gurdasani, Deepti, Tommy Carstensen, Fasil Tekola-Ayele, Luca Pagani, Ioanna Tachmazidou, Konstantinos Hatzikotoulas, Savita Karthikeyan, et al. 2015. "The African Genome Variation Project Shapes Medical Genetics in Africa." *Nature* 517(7534):327–32.
- Hamer, Dean. 2000. "Beware the Chopsticks Gene." *Molecular Psychiatry* 5(1):11–13.
- Hansen, Randall, and Desmond King. 2013. *Sterilized by the State*. Cambridge, UK: Cambridge University Press.
- Harden, Paige, Benjamin Domingue, Daniel Belsky, Jason Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot Tucker-Drob, and Kathleen Mullan Harris. 2020. "Genetic Associations with Mathematics Tracking and Persistence in Secondary School." *NPJ Science of Learning* 5:1. doi:10.1038/s41539-020-0060-2.
- Herd, Pamela. 2010. "Education and Health in Late Life among High School Graduates." *Journal of Health and Social Behavior* 51(4):478–96.
- Herd, Pamela, Deborah Carr, and Carol Roan. 2014. "Cohort Profile: Wisconsin Longitudinal Study." *International Journal of Epidemiology* 43(1):34–41.
- Herd, Pamela, Jeremy Freese, Kamil Sicinski, Benjamin Domingue, Kathleen Mullan Harris, Caiping Wei, and Robert Hauser. 2019. "Genes, Gender Inequality, and Educational Attainment." *American Sociological Review* 84(6):1069–98.
- Herd, Pamela, Brian Goesling, and James House. 2007. "Socioeconomic Position and Health." *Journal of Health and Social Behavior* 48(3):223–38.
- Holmes, Ian. 2018. "What Happens When Geneticists Talk Sloppily about Race." *The Atlantic*, April 25. <https://www.theatlantic.com/science/archive/2018/04/reich-genetics-racism/558818/>.
- House, James, Paula Lantz, and Pamela Herd. 2005. "Continuity and Change in the Social Stratification of Aging and Health over the Life Course." *The Journals of Gerontology: Social Sciences* 60(2):S15–26.
- House, James, James Lepkowski, Ann Kinney, Richard Mero, Ronald Kessler, and Regula Herzog. 1994. "The Social Stratification of Aging and Health." *Journal of Health and Social Behavior* 35:213–34.
- Jackson, Kirabo. 2018. "Does School Spending Matter? The New Literature on an Old Question." *National Bureau of Economic Research, Working Paper Series, Working Paper 25368*. <https://www.nber.org/papers/w25368>.
- Johnson, Rebecca, Ramina Sotoudeh, and Dalton Conley. 2020. "Polygenic Scores for Plasticity." *bioRxiv*. doi:10.1101/2020.08.30.274530.
- Kahn, Jonathan, Alondra Nelson, Joseph Graves, Sarah Abel, Ruha Benjamin, Sarah Blacker, Catherine Bliss, et al. 2018. "How Not to Talk about Race and Genetics." *Buzzfeed News*, March 30. <https://www.buzzfeednews.com/article/bfopinion/race-genetics-david-reich>.
- Khera, Amiti, Mark Chaffin, Krishna Aragam, Mary Haas, Carolina Roselli, Seung Hoan Choi, Pradeep Natarajan, et al. 2018. "Genome-Wide Polygenic Scores for Common Diseases Identify Individuals with Risk Equivalent to Monogenic Mutations." *Nature Genetics* 50(9):1219–24.
- Kim, Michelle, Kane Patel, Andrew Teng, Ali Berens, and Joseph Lachance. 2018. "Genetic Disease Risks Can Be Misestimated across Global Populations." *Genome Biology* 19(1):1–14.
- Lachance, Joseph, and Sarah Tishkoff. 2013. "SNP Ascertainment Bias in Population Genetic Analyses: Why It Is Important, and How to Correct It." *Bioessays* 35(9):780–86.
- Lambert, Jonathan. 2019. "Human Genomics Research Has a Diversity Problem." *NPR*, March 21. <https://www.npr.org/sections/health-shots/2019/03/21/705460986/human-genomics-research-has-a-diversity-problem>.
- Lee, Catherine. 2009. "'Race' and 'Ethnicity' in Biomedical Research?" *Social Science and Medicine* 68(6):1183–90.
- Lee, James, Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghizian, Meghan Zacher, Tuan Anh Nguyen-Viet, et al. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50(8): 1112–21.
- Leonard, Thomas. 2017. *Illiberal Reformers*. Princeton, NJ: Princeton University Press.
- Link, Bruce, and Jo Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* 35(5):80–94.
- Liu, Hexuan. 2018. "Social and Genetic Pathways in Multigenerational Transmission of Educational Attainment." *American Sociological Review* 83(2):278–304.
- Liu, Hexuan, and Guang Guo. 2015. "Lifetime Socioeconomic Status, Historical Context, and Genetic Inheritance in Shaping Body Mass in Middle and Late Adulthood." *American Sociological Review* 80(4):705–37.
- Lovejoy, Paul. 2011. *Transformations in Slavery: A History of Slavery in Africa*. Cambridge, UK: Cambridge University Press.
- Martschenko, Daphne, Sam Trejo, and Benjamin Domingue. 2019. "Genetics and Education: Recent Developments in the Context of an Ugly History and an Uncertain Future." *AERA Open* 5(1). doi:10.1177/2332858418810516.
- Mathieson, Iain, and Aylwyn Scally. 2020. "What Is Ancestry?" *PLoS Genetics* 16(3):e1008624.
- McQueen, Matthew, Jason Boardman, Benjamin Domingue, Andrew Smolen, Joyce Tabor, Ley Killea-Jones, Carolyn Halpern, Eric Whitsel, and Kathleen Mullan Harris. 2015. "The Add Health Sibling Pairs Genome-Wide Data." *Behavior Genetics* 45(1):12–23.

- Merrick, Jane. 2013. "Michael Gove Held Talks with 'IQ Genes Professor.'" *The Independent*, October 12. <https://www.independent.co.uk/news/education/education-news/michael-gove-held-talks-with-iq-genes-professor-8876500.html>.
- Mersha, Tesfaye, and Tilahun Abebe. 2015. "Self-Reported Race/Ethnicity in the Age of Genomic Research." *Human Genomics* 9(1):1–15.
- Mills, Melinda, Nicola Barban, and Felix Tropf. 2020. *An Introduction to Statistical Genetic Data Analysis*. Boston, MA: MIT Press.
- Mills, Melinda, and Charles Rahal. 2019. "A Scientometric Review of Genome-Wide Association Studies." *Communications Biology* 2(1):1–11.
- Mills, Melinda, and Charles Rahal. 2020. "The GWAS Diversity Monitor Tracks Diversity by Disease in Real Time." *Nature Genetics* 52(3):242–43.
- Mills, Melinda, and Felix Tropf. 2020. "Sociology, Genetics, and the Coming of Age of Sociogenomics." *Annual Review of Sociology* 46:553–81.
- Morning, Ann. 2011. *The Nature of Race*. Berkeley, CA: University of California Press.
- Morning, Ann, Hannah Brückner, and Alondra Nelson. 2019. "Socially Desirable Reporting and the Expression of Biological Concepts of Race." *Du Bois Review: Social Science Research on Race* 16(2):439–55.
- Morning, Ann. 2014. "Does Genomics Challenge the Social Construction of Race?" *Sociological Theory* 32(3):189–207.
- Morris, Tim, Neil Davies, and George Davey Smith. 2020. "Can Education Be Personalised Using Pupils' Genetic Data?" *Elife* 9(2020):e49962.
- Murray, Charles. 2020. *Human Diversity: Biology of Gender, Race, and Class*. New York, NY: Twelve.
- Nielsen, Rasmus, Joshua M. Akey, Mattias Jakobsson, Jonathan K. Pritchard, Sarah Tishkoff, and Eske Willerslev. 2017. "Tracing the Peopling of the World through Genomics." *Nature* 541(7637):302–10.
- Nolte, Ilja Maria, Felix Christian Tropf, and Harold Snieder. 2019. "Missing Heritability of Complex Traits and Diseases." *eLS*. doi:10.1002/9780470015902.a0028223.
- Novembre, John, Toby Johnson, Katarzyna Bryc, Zoltán Kutalik, Adam R. Boyko, Adam Auton, Amit Indap, et al. 2008. "Genes Mirror Geography within Europe." *Nature* 456:98–101.
- Oleksyk, Taras K., Vladimir Brukhin, and Stephen J. O'Brien. 2015. "The Genome Russia Project: Closing the Largest Remaining Omission on the World Genome Map." *GigaScience* 4(1):s13742–015.
- Panofsky, Aaron, and Catherine Bliss. 2017. "Ambiguity and Scientific Authority." *American Sociological Review* 82(1):59–87.
- Panofsky, Aaron, and Joan Donovan. 2019. "Genetic Ancestry Testing among White Nationalists." *Social Studies of Science* 49(5):653–81.
- Peterson, Roseann, Karoline Kuchenbaecker, Raymond Walters, Chia-Yen Chen, Alice Popejoy, Sathish Periyasamy, Max Lam, et al. 2019. "Genome-Wide Association Studies in Ancestrally Diverse Populations." *Cell* 179(3):589–603.
- Pew Research. 2018. "Key Facts about Black Immigrants in the US." <https://www.pewresearch.org/fact-tank/2018/01/24/key-facts-about-black-immigrants-in-the-u-s/>.
- Phelan, Jo, Bruce Link, and Naumi Feldman. 2013. "The Genomic Revolution and Beliefs about Essential Racial Differences." *American Sociological Review* 78(2):167–91.
- Phelps, Steven, and Robbee Wedow. 2019. "What Genetics Is Teaching Us about Sexuality." *New York Times*, August 29. <https://www.nytimes.com/2019/08/29/opinion/genetics-sexual-orientation-study.html>.
- Plomin, Robert. 2018. *Blueprint: How DNA Makes Us Who We Are*. Boston, MA: MIT Press.
- Plomin, Robert. 2019. "Talk for the UK Behavioral Insights Team." October 8. Video, 1:00:02. [https://www.youtube.com/watch?v=LZ9lu7\\_MpwQ](https://www.youtube.com/watch?v=LZ9lu7_MpwQ).
- Plomin, Robert, and Sophie von Stumm. 2018. "The New Genetics of Intelligence." *Nature Reviews Genetics* 19(3):148–59.
- Reich, David. 2018a. "How Genetics Is Changing Our Understanding of Race." *New York Times*, March 23. <https://www.nytimes.com/2018/03/23/opinion/sunday/genetics-race.html>.
- Reich, David. 2018b. "How to Talk about 'Race' and Genetics." *New York Times*, March 30. <https://www.nytimes.com/2018/03/30/opinion/race-genetics.html>.
- Reich, David. 2019. *Who We Are and How We Got Here*. Oxford, UK: Oxford University Press.
- Roberts, Dorothy. 2015. "Can Research on the Genetics of Intelligence Be 'Socially Neutral'?" *Hastings Center Report* 45(S1):S50–53.
- Roberts, Dorothy. 2011. *Fatal Invention*. New York, NY: New Press/ORIM.
- Roberts, Dorothy (DorothyERoberts). 2020. "Fatal Invention spells out that race is not a real BIOLOGICAL creation, but it is a real POLITICAL system—invented to enforce racism, in part by claiming Black people are naturally less than human. #BlackLivesMatter directly contests that racist ideology. Perfectly consistent." Twitter, June 27, 9:01 a.m. <https://twitter.com/dorothyroberts/status/127687822232297474?lang=en>.
- Roberts, Dorothy, and Oliver Rollins. 2020. "Why Sociology Matters to Race and Biosocial Science." *Annual Review of Sociology* 46:195–214.
- Saini, Angela. 2019. *Superior: The Return of Race Science*. Boston, MA: Beacon Press.
- Saletan, William. 2018. "Stop Talking about Race and IQ." *Slate*, April 27. <https://slate.com/news-and-politics/2018/04/stop-talking-about-race-and-iq-take-it-from-someone-who-did.html>.
- Schmitz, Lauren, and Dalton Conley. 2017. "The Effect of Vietnam-Era Conscript and Genetic Potential for Educational Attainment on Schooling Outcomes." *Economics of Education Review* 61:85–97.
- Selzam, Saskia, Stuart Ritchie, Jean-Baptiste Pingault, Chandra Reynolds, Paul O'Reilly, and Robert



- Plomin. 2019. "Comparing Within- and Between-Family Polygenic Score Prediction." *The American Journal of Human Genetics* 105(2):351–63.
- Silventoinen, Karri, Aline Jelenkovic, Reijo Sund, Antti Latvala, Chika Honda, Fujio Inui, Rie Tomizawa, et al. 2020. "Genetic and Environmental Variation in Educational Attainment." *Scientific Reports* 10(1):1–11.
- Skoglund, Pontus, and Iain Mathieson. 2018. "Ancient Genomics of Modern Humans." *Annual Review of Genomics and Human Genetics* 19:381–404.
- Sohail, Mashaal, Robert Maier, Andrea Ganna, Alex Bloemendal, Alicia Martin, Michael Turchin, Charleston W. K. Chiang, et al. 2019. "Polygenic Adaptation on Height Is Overestimated." *Elife* 8:e39702. doi:10.7554/eLife.39702.001
- Sonnega, Amanda, Jessica Faul, Mary Beth Ofstedal, Kenneth M. Langa, John Phillips, and David Weir. 2014. "Cohort Profile: The HRS." *International Journal of Epidemiology* 43(2):576–85.
- Sudmant, Peter, Tobias Rausch, Eugene Gardner, Robert Handsaker, Alexej Abyzov, John Huddleston, Yan Zhang, et al. 2015. "An Integrated Map of Structural Variation in 2,504 Human Genomes." *Nature* 526(7571):75–81.
- Tillery, Alvin. 2018. "The First-Ever National Survey on Race and Genetics Finds Americans Have a Nuanced Understanding of Race." Northwestern University: Center for the Study of Diversity and Democracy. [https://csdd.northwestern.edu/documents/csdd-survey-results\\_23andme-poll\\_august-20181.pdf](https://csdd.northwestern.edu/documents/csdd-survey-results_23andme-poll_august-20181.pdf).
- Trejo, Sam, Dan Belsky, Jason Boardman, Jeremy Freese, Kathleen Mullan Harris, Pamela Herd, Kamil Sicinski, and Ben Domingue. 2018. "Schools as Moderators of Genetic Associations with Life Course Attainments." *Sociological Science* 2(5):513–40.
- Tropf, Felix, Hong Lee, Renske M. Verweij, Gert Stulp, Peter Van Der Most, Ronald De Vlaming, Andrew Bakshi, et al. 2017. "Hidden Heritability due to Heterogeneity across Seven Populations." *Nature Human Behaviour* 1(10):757–65.
- Tsosie, Krystal, Joseph Yracheta, and Donna Dickenson. 2019. "Overvaluing Individual Consent Ignores Risks to Tribal Participants." *Nature Reviews Genetics* 20(9):497–98.
- Turkheimer, Eric, Kathryn Paige Harden, and Richard Nisbett. 2017. "Charles Murray Is Once Again Peddling Junk Science about Race and IQ." *Vox*, May 18. [https://www.vox.com/the-big-idea/2017/5/18/1565563\\_8/charles-murray-race-iq-sam-harris-science-free-speech](https://www.vox.com/the-big-idea/2017/5/18/1565563_8/charles-murray-race-iq-sam-harris-science-free-speech).
- Ujma, Péter, Nóra Eszlári, András Millinghoffer, Bence Bruncsics, Péter Petschner, Péter Antal, Bill Deakin, György Bagdy, and Gabriella Juhász. 2020. "Educational Attainment Polygenic Scores in Hungary." *bioRxiv*. doi:10.1101/2020.01.13.905034v2.
- Vyas, Darshali, Leo Eisenstein, and David Jones. 2020. "Hidden in Plain Sight." *New England Journal of Medicine* 383(9):874–82.
- Wailoo, Keith, Alondra Nelson, and Catherine Lee. 2012. *Genetics and the Unsettled Past*. Newark, NJ: Rutgers University Press.
- Warren, John Robert, Jennifer Sheridan, and Robert Hauser. 2002. "Occupational Stratification across the Life Course." *American Sociological Review* 67(3):432–55.
- Wedow, Robbee, Meghan Zacher, Brooke Huibregtse, Kathleen Mullan Harris, Benjamin Domingue, and Jason Boardman. 2018. "Education, Smoking, and Cohort Change." *American Sociological Review* 83(4):802–32.
- Williams, David, Jourdyn Lawrence, and Brigitte Davis. 2019. "Racism and Health." *Annual Review of Public Health* 40:105–25.

## AUTHOR BIOGRAPHIES

**Pamela Herd** is a professor of public policy at the McCourt School of Public Policy at Georgetown University. Her research is at the intersection of inequality, aging, policy, and biodemography. Her recent work on health has focused on how social inequalities and biological processes interact with each other to shape inequality over the life course. Her work is funded by the National Institutes for Health, National Institute on Aging, and the Russell Sage Foundation.

**Melinda C. Mills**, MBE, FBA, is director of the Leverhulme Centre for Demographic Science, University of Oxford, where she is also the Nuffield Professor of Sociology. Her research spans multiple topics in demography, empirical sociology, statistics, and genetics. Her recent work focuses on sociogenomics, combining a social science and molecular genetic approach to the study of behavioral outcomes with a focus on reproduction (fertility), chronotype, nonstandard, precarious employment, and assortative mating.

**Jennifer Beam Dowd** is an associate professor of demography and population health at the University of Oxford, where she is also deputy director of the Leverhulme Centre for Demographic Science. She is a quantitative social and health scientist with interdisciplinary training in demography, economics, epidemiology, and infectious disease. Her research seeks to understand how social and biological processes interact over the life course and how social factors "get under the skin" to impact health.