Zelizer, Viviana A. 1994. *The Social Meaning of Money*. New York: BasicBooks.

Zelizer, Viviana A. 2010. ''Chuck Tilly and Mozart.'' *The American Sociologist* 41(4):423–8.

Zelizer, Viviana A., and Charles Tilly. 2006. ''Relations and Categories.'' *Psychology of Learning and Motivation* 47:1–31.

The Arrival of Social Science Genomics

Jeremy Freese
Stanford University
jfreese@stanford.edu

''The genetics revolution may be well underway,'' write Dalton Conley and Jason Fletcher in *The Genome Factor*, ''but the social genomics revolution is just getting started'' (p. 11). They are not alone in their excitement for recent developments bringing together social science and genetic research. Decades from now, folks may well look back at this time as the start of a golden age for the field.

Of course, declaring the rise of social genomics is bound to make many sociologists nervous. Sociology often identifies itself as a style of explanation, with references to genetic differences considered the quintessential ''Not Sociology'' explanation. References to genetic differences are also regarded, with abundant historical justification, as dangerous for many of the moral and ideological commitments that sociologists overwhelmingly share. Sociologists at the intersection of critical race theory and science studies have been particularly vocal opponents of any role for genomic data in social science. A recent book by Catherine Bliss, *Social by Nature*, contends that social science genomics needs a ''wake-up call'' about the naïveté of its approach and its implication in social harm.

Conley and Fletcher provide a commendable overview to social science genomics and anticipate some of the criticisms that Bliss raises. But *The Genome Factor* is written for a broad audience, and here I wish to discuss social science genomics in a way directed more specifically to sociologists. I will focus especially on what I regard as the most promising development for social scientists: polygenic scores. I will talk first about the science of polygenic scores and then about some of the policy and moral questions at hand, offering pointers to *The Genome Factor*, *Social by Nature*, and several other books along the way.

Why is now different?

Social science genomics has so far unfolded like a three-act play. Act I relied on the statistical model of data on twins, adoptees, and other familial relationships. The fundamental empirical product was a *heritability* estimate, intended as the proportion of variation in an outcome among members of a population that is due, in one way or another, to genetic differences. This paradigm generated an enormous number of studies cataloging substantial genetic influence nearly everywhere it looked. The same tools that quantified the heritability of height and schizophrenia also showed the heritability of abilities, attitudes, experiences, and attainments.

Act I proceeded without direct measures of genes themselves. This indirectness was one reason the field struggled to overcome its skeptics, who could latch onto one or another assumption as grounds for doubts about the enterprise. Worse, applying these tools to more complex questions about genetic influence required adding yet more assumptions, making already indirect inferences even more so.

Act II opens with molecular data on genetic differences becoming available. This phase began with researchers having data only on a very small number of sites on the genome, for only a few dozen or hundred—later, a few thousand—participants. The logic of analysis was otherwise familiar—more familiar than for the variance decomposition methods that dominated Act I. Papers posed hypotheses about

the relationship between a specific genetic difference and an outcome and then presented results using a regression-type approach in which environmental measures were deployed as controls or as interaction effects.

The core problem was that there were so many potential hypotheses and so many ways of specifying measures, subgroups, and models when testing them. Statistically significant results were easy to find and publish, but findings would not replicate in other samples. Act II is now referred to as the "candidate gene era," often with an airily dismissive tone suggestive of an old friend waving off lost years of their twenties when they were in a cult. The overwhelming majority of published "discoveries" from this era are nowadays recognized as having mistaken noise for signal.

What saved the enterprise is that assaying kept getting cheaper and better. The basic unit of genomic variation is the single nucleotide polymorphism (SNP, pronounced "snip"), which is any site on the three-billion base pair genome in which many people vary. Genome-wide association studies (GWAS) became possible, involving assays of hundreds of thousands or a million or more SNPs per person. Now you can sequence the entire genome with reasonable quality for less cost per person than what it took to assay a single SNP two decades ago.

Act III began by disposing with conventional hypothesis testing in favor of a more basic brute-force approach. Classic GWAS evaluates the association between every assayed genetic variant and some outcome ("phenotype"), but using an extremely stringent standard of statistical significance ($p < .00000005$). To maximize sample size, researchers combined many different samples into consortia.

GWAS research is mostly focused on health. But, as in Act I, the methods started being applied more broadly, including to the study of educational attainment, the most studied non-demographic variable in sociology. Educational attainment was a great fit for GWAS because information about it is routinely collected in medical studies. Nevertheless, to many, it still seemed like a longshot.

The Social Science Genetics Consortium (SSGAC) published its first results for educational attainment in 2013. Using a combined sample size of over 100,000, it found 3 "hits"—significant associations between a SNP and phenotype by the stringent standard of GWAS. In 2016, SSGAC published a follow-up with over 300,000 respondents, finding 71 significant hits. Recently, SSGAC has presented work involving more than a million respondents and finding more than a thousand hits.

To be clear: nobody is saying genes *determine* educational attainment. Nobody is even saying genetic information *predicts* educational attainment all that well. But standard sociological variables do not predict educational attainment that well either. The SSGAC study's results indicate that genomic information presently can be combined to predict a person's educational attainment about as well as one can from knowing the educational attainment of one of that person's parents (Lee, forthcoming).

Conley and Fletcher describe GWAS as the cornerstone of what they call "small effects" genetics. The idea is that even when a phenotype's heritability is substantial, genetic influence predominantly arises from a very large number of very small effects of genetic differences spread throughout the genome. This characterizes how genetic influence works for many highly heritable complex diseases, like heart disease (Nikpay et al. 2015), and it is now clear the same is true for educational attainment.

What are polygenic scores?

GWAS studies from large consortia post their results online. These are basically a few million regression coefficients, one for each SNP. Others can then use the GWAS results to generate predictive scores for individuals *in other samples* with genomic data. These are called *polygenic scores*. Polygenic scores summarize all the genetic associations captured in the GWAS. Being derived from DNA, polygenic scores are intrinsically prospective: they are not affected by anything that happens to individuals during their lives.

Given the false starts of the candidate gene era, the most important fact about polygenic scores is that they *work*. The best evidence comes from siblings. If studies only involved

samples of unrelated individuals, then a correlation between polygenic scores and outcomes could just be picking up information about ancestry. An often-used illustration is that genetic differences are surely correlated with using chopsticks when eating, just because of statistical genetic differences between those of East Asian ancestry versus others.

Full siblings, however, have exactly the same ancestors. (With genetic data, you can tell whether respondents really are full siblings regardless of what their self-reports might say.) Full siblings are natural experiments: genetic differences between them are random. Yet evidence from multiple samples indicates that, when siblings differ in educational attainment, the sibling with the higher polygenic score is more likely to be the one with more education. These results seem inexplicable except by the polygenic score capturing *causal* influences of genetic differences on educational attainment.

Why should sociologists care?

Even sociologists who study individual-level outcomes often reason that, because they study "social causes" and genetic differences are not social causes, genes are somebody else's problem. The tenability of this sort of pragmatic partitioning is likely undermined by the complex causal interdependence of human differences. More importantly, polygenic scores are not simply a means of detecting genetic influence. They have all kinds of possibilities for advancing social science inquiry in ways useful to social scientists who do not care about genes per se.

1) Polygenic scores pose a *puzzle*. Polygenic scores are purely predictive scores, like credit scores. So imagine having a dataset of individuals' credit scores, knowing those scores summarized risk of credit default, but not knowing how or why they worked to do so. We treat this as a puzzle and try to reverse-engineer which individual characteristics were associated with the credit score; in so doing, we stand to gain insights into the causes of default. This is where we are with educational attainment: genes do not directly cause educational attainment,

of course, but they do affect it somehow by influencing chains of intermediate traits, behaviors, and environments. Pursuing why the scores predict educational differences provides another perspective into understanding the causes of educational attainment.

For example, existing social science literature contains many references to "noncognitive skills." The vagueness of this term might itself suggest the value of social scientists in this area taking analytic help wherever they can get it. Indeed, many characteristics regularly mentioned as "noncognitive differences" are notoriously difficult to measure at scale, and observed traits also reflect environments in ways that confound estimates of their causal influence. If one grants the evidence that these traits are *partly* genetically influenced, then polygenic scores give some leverage for understanding effects of these traits on educational attainment, without any assumptions that traits are only determined by genes or that they cannot be modified by interventions.

2) Polygenic scores also offer obvious value as a "*control variable*." Social scientists are seldom asked directly to address potential confounding of results by genetic differences. But various watchwords of rising standards for causal inference—"unobserved heterogeneity," "endogeneity bias," "selection bias"—index issues that may be partly addressed by information about genes. The premier U.S. study of the demography of aging, the Health and Retirement Study, has begun making polygenic scores for various phenotypes it constructs available as a user product, and other large datasets with genomic data are working to follow suit. Scores reflect the labors of a pipeline of cutting-edge techniques that is longer than what any one individual could master, but the result is that polygenic scores can be included in regression models as readily as more familiar survey-based measures.

3) Polygenic scores may serve as *moderators* (aka "interaction effects"). That is, scores can help us better understand how genetic differences modify effects of environmental causes. For example, polygenic scores for smoking appear predominantly to reflect neurophysiological differences in how different people's bodies react to nicotine. In

the name of promoting public health, many states levy a high—and highly regressive—"sin tax" on purchases to financially punish those who smoke. One of the examples Conley and Fletcher discuss in *The Genome Factor* provides evidence that polygenic influences are implicated in whose smoking is deterred by higher taxes, which may raise moral questions about the fairness of these taxes.

The prospects of identifying gene-environment interactions are arguably the biggest driver for social science interest in genetics. They raise a fundamental question for social scientists: how does one imagine that social-environmental causes affect people? Evidence is clear that similar experiences affect different people differently, and this heterogeneity far exceeds anything that conventional social science measures can explain. Is the answer just the old joke about the first law of sociology: "some do, and some don't"? If one grants the excruciatingly obvious point that psychological and physiological differences are part of why the same causes affect different people differently, then the partial influence of genes on such differences seems poised to provide a valuable lever toward better understanding the dynamic interactions among mind, body, and social world.

4) Polygenic scores may help us understand *social processes of life course divergence* more broadly. One of the most misleading bits of pop social science to penetrate public consciousness in recent years is the idea, most closely associated with Malcolm Gladwell, that all that separates everyone from approximately equal eliteness at anything is 10,000 hours of deliberate practice. Instead, across countless domains of performance, evidence makes plain that the same amount of investment yields different returns for different people. Social incentives and processes strongly push people to invest in domains they show aptitude for. When aptitude and performance are correlated—for example, when kids who are already good at reading read more than kids who aren't—differences can widen. To the extent those early displays of aptitude are partly genetically influenced, genetic differences can set off a chain of environmental differences that serve to expand and stabilize differences.

Polygenic scores offer possible leverage into processes of cumulative advantage and disadvantage that sociologists have long known pervade human life. The scores also provide a way of documenting how some environments offer different returns for people with similar underlying scores. A leading idea here is that social advantages increase children's ability to realize their "potential," which polygenic scores can help interrogate, specifically in identifying the aspects of environment that matter most.

But isn't this all just genetic determinism?

The point of the above examples is that the big payoff to social science from polygenic scores is better understanding of the true influences of social environments. I have emphasized that polygenic scores only explain a modest portion of complex outcomes and, even then, much of that may involve ways that environments amplify effects of genetic differences. Nevertheless, a fact of the sociological life is that you cannot say that genes influence behavior in any sort of specific way without folks accusing you either of espousing genetic determinism or—in their minds, more charitably—of unwittingly encouraging a rhetoric of genetic determinism.

Sure enough, in *Social by Nature*, Bliss accuses social science genomics of "reproducing genetic determinism" (e.g., p. 61), and that charge is central to the variegated harms she sees as ensuing from the enterprise. Her most sustained critique to this end is that social science genomics pushes a "gene-first" approach that "filter[s] everything through a gene-gene prism" and relegates environments to the causal backbench.

Making this critique requires Bliss to overrule many examples in which members of the field profess their close interest in environments and deny genetic determinism. Sometimes she does this by erroneous assertion, as when she announces the absurd and easily-refuted-by-Google claim that social science genomics focuses on "finding gene-gene interactions more than gene-environment ones" (p. 61). Sometimes she misrepresents work, as when she criticizes a paper by Daw et al. (2013) for ignoring gene-

environment interaction in favor of gene-gene analysis, even though the paper does not in fact include any gene-gene analysis and presents itself as pursuing environmental moderation of genetic effects. Sometimes she misunderstands biological concepts, as when she represents the study of one phenomenon (pleiotropy) as a ''gene-gene analysis'' when it is approximately the opposite of gene-gene analysis (p. 77). Sometimes she misunderstands social science, as when she presents utterly conventional ways of measuring educational attainment as ''new metrics'' that social science genomics has devised to ''translate social measures into genetic ones'' (p. 74).

Bliss also makes her critique using various quasi-anonymized interviews with researchers in the field—I am Interview #9—but these are also replete with mistakes and dubious interpretations. Bliss evinces all sorts of misunderstandings about the biographies, intellectual orientations, and relationships among the members of the research community that she criticizes. (I count ten facts presented about me, for example, and six are wrong.) Indeed, Bliss's interview quotes contain transcription howlers that raise doubts about how well she understood what researchers were saying. A researcher obviously talking about null hypothesis testing is reported as saying ''no hypothesis testing'' (p. 125); a reference to estimating systems of equations becomes the gibberish ''estimate existence of equations'' (p. 121); and an obvious reference to ''sexual partners'' is misunderstood as ''factorial partners'' (p. 68).

I think any academic would be irked by such rank sloppiness in a book about one's own field. But this is a book that paints social science genomics in such a negative light that the headline of *Nature*'s review alludes to Nazism (Comfort 2018). The serial misinterpretations in *Social by Nature* ultimately service the ethnographer-savior-complex for which critical science studies has unending appetite—the narrative of a society threatened by hard-charging scientists morally blind to issues that only a circle of qualitative researchers have sufficient virtue to see clearly. When Bliss depicts those who do social science genomics as convinced that ''genetics [will] direct us to a more

utopian society''—or when she claims we depict educational attainment as ''no different from Alzheimer's or autism''—the point is not accurate description but to cast us as the maniacal villains of an intellectual melodrama in which the ethnographer is the moral hero. She declares toward the end that genetic science ''will certainly set us back centuries in terms of stereotypes and social relations, and propel us forward into a dystopian future.''

To return to the question with which I started this section, then: no, this is not at all genetic determinism, no matter how often or dramatically critical sociology announces otherwise.

What about epigenetics?

Polygenic scores are based on DNA, which, with limited exceptions, is the same in every cell in the body and does not change over people's lives. Polygenic scores thus do not depend on when or how a person's DNA was collected. In contrast, many social scientists have heard about ''epigenetics,'' which covers processes that modify the expression of genes in ways that do differ across cells and over time. The most prominent example is DNA methylation, in which methyl groups attach to a site on the DNA molecule, which usually prevents that part of the molecule from being transcribed into RNA, rendering it inactive.

As Conley and Fletcher note in *The Genome Factor*, ''social scientists have jumped on the epigenetic bandwagon in droves'' (p. 214). Social scientists may be disappointed to find that the *book* only really engages epigenetics as an appendix. An enthusiastic treatment of the possibilities of epigenetics for understanding human behavior is David Moore's *The Developing Genome*.

The work that has most captured public imagination about epigenetics—recounted in detail by Moore—involves experimentally rearranging rat litters. Baby rats were randomly reassigned to mothers who varied in how much they licked and groomed their pups. The pups who were licked and groomed less grew up to be adults who responded more strongly to mild rat stressors, like being plopped into an unfamiliar environment. Subsequent work connected

licking and grooming to DNA methylation at sites associated with brain development.

Epigenetics provides paths for environmental circumstances to moderate genetic causation at the cellular level. It thereby affords, for example, a possible way for stressful life events to get "under the skin" and affect biological outcomes. Also, given that much methylation happens very early in life, epigenetics has been presented as providing biological evidence for bolstering early childhood interventions.

Intrigue about epigenetics also comes from indications of methylation patterns being transmitted from mothers to offspring. That this can happen is not biologically obvious, as methylation patterns are largely wiped clean in the process of reproduction. But, for example, if the licked-and-groomed pups grow up to do more licking and grooming as parents, then the same patterns of methylation they underwent as pups might also be observed in their own pups, through a process reminiscent of one form of "social reproduction." More provocative are more direct and potentially enduring mechanisms by which methylation patterns are reproduced over generations, which harken to the ideas of acquired characteristics of parents being biologically inherited by offspring, ideas most famously associated with Lamarck.

Lamarckian inheritance, incidentally, is often currently construed as more consistent with left-leaning politics than standard Darwinian inheritance. Maybe this is sympathy for the underdog, although Darwinism remains a bit of an underdog itself in the United States considering how many folks don't believe in it. But Lamarckism is certainly not *intrinsically* leftist. One of its most important early champions was Herbert Spencer, for whom Lamarckism implied a natural progress that would only be obstructed by humanitarian state interventions. Lamarckism's most famous association with leftist politics is Trofim Lysenko's ascension in Soviet agriculture, which led to the punishment and death of many orthodox Soviet geneticists and worse Soviet crop yields. Nevertheless, recent enthusiasm for epigenetics has led to some efforts to rehabilitate Lysenko's reputation. Historian Loren Graham's *Lysenko's Ghost* takes up the question and concludes "where [Lysenko] was right, he was not original; where he was original, he was not right" (p. 141).

In any event, how valuable epigenetics will ultimately prove for social scientists is currently hard to discern. I am personally befuddled by how readily some folks accept mother rats' licking and grooming as homologous with suburban helicopter parents. Clearer-eyed assessments of epigenetics' prospects may only emerge after some of its hype has dissipated.

Until then, readers are urged to be wary of being overly swayed by trade book accounts. Take *The Telomere Effect*, by Nobel Prize-winning biologist Elizabeth Blackburn and health psychologist Elissa Epel. Telomeres are repetitive DNA at the ends of chromosomes that protect the strands from deterioration, kind of like the little caps at the end of shoelaces. Telomeres shorten as cells divide; and, although they can be renewed in other ways, average telomere length shortens with age.

As "effect" in the title suggests, the book is about the enormous causal power of telomeres in aging. If you "protect your telomeres"—Telomere Tips are provided at the end of each chapter—then you will slow your aging rate. The book's health recommendations are ultimately unobjectionable, emphasizing the same combination of diet, exercise, sleep, and stress reduction recommended by endless other books that offer the same advice without ever mentioning telomeres. The book also makes pleasing shout-outs to social desiderata like reducing discrimination and inequality.

Even so, *The Telomere Effect* serially overstates the evidence for the causal effects of telomere length. Sure, it often uses language that academics will recognize as synonyms for "association," such as "predicts," "is linked to," or "marker," but the book does nothing to disabuse less sophisticated readers of the difference. Such a reader would never suspect, for example, that studies have repeatedly found that black Americans have longer average telomere length than white Americans, despite blacks having shorter life expectancies and less exposure to literally dozens of things the book identifies as good for one's telomeres. These are fronts on which both *The Genome Factor* and *The*

Developing Genome deserve praise: even though the books are written for a popular audience, they present information in a way that does not make the readers who have knowledge of the underlying science feel dirty.

What about race?

The biggest reason discussions of genetic differences discomfit social scientists is, of course, race. The use of dubious biological arguments to justify racial inequalities has a long history, well documented by others (for example, law professor Dorothy Roberts's *Fatal Invention* [2011]). Recent examples are also not hard to find. For example, in *A Troublesome Inheritance*, former *New York Times* journalist Nicholas Wade argues that differences between more and less developed societies emerge from genetic differences in persons that are amplified into large societal differences.

Wade's book is replete with problems, which have been well-cataloged by others (e.g., Dobbs 2014). My reason for mentioning *A Troublesome Inheritance* here is to note how pervasively it relies on the politicization of social science to advance an argument that has all sorts of factual and logical problems. Why should the reader believe a lone science journalist when that journalist is making arguments that most relevant experts would dispute? Wade's answer, asserted throughout, is that social science expertise is untrustworthy because of the ''creed'' that social scientists must adhere to when they talk about race. Early on, Wade quotes an American Sociological Association statement about race and then writes that ''social scientists' official view of race is designed to support the political view that genetics cannot possibly be the reason why human societies differ'' (p. 5).

Wade is far from the only author to draw from that well. Everybody knows that everybody knows some statements about race are more ''politically correct'' than others. This, in turn, provides a rhetorical weapon against experts: either they are too ideologically blind to see the truth, or they are too cowed by politics to be able to tell the truth. It is the popular science counterpart of a blowhard bellowing ''FAKE NEWS.''

To their credit, in *The Genome Factor*, Conley and Fletcher refrain from repeating common, politically expedient arguments that experts know full well are misleading. For instance, all sorts of ugly arguments about race and genetics could be immediately dismissed if it were true that meaningful genetic change required longer timescales than the continental separation of human populations. In *A Troublesome Inheritance*, Wade deploys the fact that this is incorrect—differences certainly can emerge in a small number of generations—as his first salvo in positioning himself as the one willing to announce truths that social scientists are too squeamish to acknowledge. Conley and Fletcher, meanwhile, explain the facts forthrightly and then explain why they believe that these facts do not have the broad implications some have ascribed to them.

As for the most uncomfortable question of all—whether group differences partly reflect genetic differences—Conley and Fletcher ultimately conclude that ''research and theory suggest that genetic differences are a potential—but highly unlikely— explanation for national, racial, or ethnic differences in behavior and socioeconomic success'' (p. 101). They spend several pages recounting reasons for that ''highly unlikely.''

Of course, even experts disagree about which assertions represent a forthright discussion of politically uncomfortable matters versus an overly credulous one. For example, while I found most of *The Genome Factor* praiseworthy, I am extremely skeptical of one economic study they present at length (Ashraf and Galor 2013), which contends that the genetic diversity of a society has implications for its historic economic development. The study suggests that too little diversity (e.g., indigenous Americas) and too much (e.g., Africa) pose hindrances, with a genetic Goldilocks zone (e.g., Europe and Asia) in the middle. The study involves small-sample, society-as-a-unit-of-analysis inference that is hard for me to take seriously in the best of circumstances, and it gets completely away from the strengths of large samples with large amounts of genomic information that I think represent the real promise and challenge of genomic data for social science.

What about the social construction of race?

One chapter of *The Genome Factor* is simply titled "Is Race Genetic?" Here the authors walk a tightrope familiar for social science genomics. They write, plainly, "Race does not stand up scientifically, period" (p. 94). By this, they mean that the categories society uses for discussing race/ethnicity are not genetic distinctions. For instance, as they point out, the reasons why white Americans can usually describe themselves using various ethnic subclassifications (e.g., Irish, Italian), while black Americans usually cannot, are historical and cultural explanations, not genetic ones.

At the same time, the "social construction of race" is sometimes deployed to hint at more radical positions—for example, to lend an impression that commonly used racial categories have very little bearing on what one observes when looking at genomic data. In the United States, genetic information statistically predicts individuals' self-identified race with imperfect, but nevertheless very high, accuracy. Obviously this fact can be wielded in the service of some ugly ideas, and one may be reminded of the remark of one mid-nineteenth-century aristocrat on first hearing of Darwin's theory: "Let us hope it is not true, but if it is, let us pray that it will not become generally known."

But these statistical differences have direct research consequences, creating an unavoidable collision between genomics research and those deeply concerned about scientific racism. In *Fatal Invention*, Roberts seems eminently reasonable when she argues that, "Instead of asking, 'How do genes work differently in different racial groups?' scientists could ask, 'How do genes work in human beings?'" For those who actually work with genomic data, however, failing to take the statistical issues here seriously means that research ostensibly directed at understanding risks to "human beings" can come to incorrect conclusions, in ways that can have perverse consequences regarding race.

To see how, we can return to *A Troublesome Inheritance*. Nicholas Wade describes some candidate gene work that links a genetic variant to violent behavior; this work was presented in "human beings" terms with minimal consideration of race. Wade highlights, correctly, that the genetic variant in question is much more common in black Americans than white Americans. He might as well draw dots labeled "genes," "crime," and "blacks" and invite readers to connect them. But, without getting into technical details, the research Wade cites as showing the association between the variant and aggressive behavior did not adequately address the potential confounding by racial/ethnic differences in gene frequencies. As a result, environmental causes of racial/ethnic differences in aggressive behavior may have been recast and reinterpreted as effects of genes, and this bad causal inference about genetic differences and aggression was redeployed by Wade to suggest that science has demonstrated that blacks are genetically more aggressive than whites.

What about race and polygenic scores?

For studies using polygenic scores, currently the most common way to address the possibility of confounding by biological ancestry is to restrict analysis to a particular ancestral group. Overwhelmingly, the best available polygenic scores are based on a GWAS of European-descended persons, and researchers use genetic information to restrict their own study's sample similarly—the resulting samples regularly consist exclusively of self-identified whites. While there has been considerable expansion of GWAS data on Asian, especially Han Chinese, populations, data on African and indigenous North American ancestry population lags far behind. This is mostly due to broad global inequalities in scientific research, although it bears noting that, twenty years ago, social science activists played a key role in stalling early efforts to amass data on global genomic diversity.

Researchers have good statistical reasons to be wary of applying polygenic scores to populations with different continental ancestry than that on which GWAS is based. These dovetail with moral concerns, for which one team of researchers has provided a striking parable regarding height (Martin et al. 2016). The researchers took GWAS

results for height, based on European-descended samples, and applied them to simulated European and West African populations based on established reference panels of gene frequencies for these populations. Startlingly, this work showed there was very little overlap in distributions: that is, nearly all West Africans would have lower polygenic scores for height than nearly all Europeans. So a naïve analyst given racially diverse data might conclude that polygenic score information was revealing the genetic basis of why West Africans are shorter than Europeans. But this cannot be right, because we know from the anthropological record the actual height distributions between these populations are not so different.

All this presents a moral thicket: given the obvious potential for trouble, how acceptable is it for research using polygenic scores to exclude blacks, Latinos, and American Indians? This is the sort of complicated moral question for which humanistic sociological thinking might help.

In *Social by Nature*, Bliss's reaction to restricting samples by ancestry is just to worry that it "reinforces a sense of essential racial difference" (p. 92) and "gives validity to unwarranted biological notions of race" (p. 92). As with most critical science studies scholarship about race and genomics, her preoccupation is with the potential cognitive consequences for how people think about race more broadly, and how these cognitive consequences might, in turn, legitimate action that harms marginalized racial and ethnic groups. That preoccupation itself follows from a separate conviction that informs much of this work: namely, that the scientific promise of genetic research has been enormously oversold. After all, if social science genomics is nothing more than hopeless hooey, every moral question about it can be reduced to whether its hooey is harmless.

For decades now, the dominant refrain of critical science studies about genetics has been "old wine in new bottles." Scholars take recent work in genetic science and recast it as another iteration of a very long arc in which bad science abets narratives of inborn superiority or deficiency. Perhaps the best examples of this argument are Stephen Jay Gould's *Mismeasure of Man*

(1981) and Troy Duster's *Backdoor to Eugenics* (1990). *Social by Nature* styles itself in a similar spirit; its cover promises "shocking parallels" between past episodes of scientific racism and social science genomics today. This probably sounds disingenuous or snarky, but I could not figure out what these parallels were supposed to be, unless it was just the charge of a "gene-first" approach that I described earlier. The discussions of historical episodes and social science genomics occur separately in *Social by Nature* with no explicit articulation of how they are meant to be connected.

In any event, historical analogies are sometimes instructive. The big problem with "old wine in new bottles" as a field's go-to argument is that it invites a presumption that old wine is all the new bottles ever contain. For example, right now there is considerable buzz around "precision medicine"—the use of genomic information to guide medical treatment decisions. While some of this will certainly prove to be just hype, it seems foolhardy to bet that it is only hype. A foreseeable possibility is that, if precision medicine starts yielding real benefits, techniques may work best for white people because of their massive overrepresentation in genetic studies. Science studies has an admirable history of highlighting ways science has operated to the detriment of marginalized groups, and it would be deeply unfortunate if its radical skepticism toward genetics led it to be slow to recognize a shifting terrain toward moral questions about groups receiving less benefit as genomic research bears fruit.

What about policy?

Anyone who has had extended interactions with health providers can intuit why precision medicine seems to be a holy grail. Treatment often involves considerable trial and error, tweaking, and decisions based on percentages, where any information that would better inform those decisions could reduce suffering and save lives. Consequently, when thinking about how genomic data may influence policy outside of health interventions, analogies to precision medicine often dominate the foreground. I think social science genomics has mostly a weak case for

its policy significance on these grounds, and instead its big promise for policy impacts is different.

One can see the problem in the book *G Is for Genes*, by behavioral geneticists Kathryn Asbury and Robert Plomin. They focus on genetics and education, based mostly on their longtime work with a large twin sample in the United Kingdom. The book is ostensibly very policy oriented. The authors state at the outset that ''It's time for educationalists and policy makers to sit down with geneticists to apply [genetic research] findings to educational practice'' (p. 3). One chapter summarizes ''Eleven Policy Ideas'' and another proposes what they would do if they were ''Education Secretary for a Day.''

Plomin has been a leading figure for years in genetics and IQ research, and so many sociologists may approach the book imagining, well, *Bell Curve II.* Readers may then be surprised by the familiarly left-leaning nature of many of the book's policy proposals: free preschool for everyone, extra support for low-SES families from birth, helping poorer children have equal access to extracurricular activities.

The biggest policy implication that Asbury and Plomin draw is that education should be more individualized, with more tailoring of how students are taught in core subjects and more flexibility in which subjects they pursue. The argument for how available genetic research justifies this conclusion largely boils down to arguing that genetic research makes clear that children differ, that they respond to environments differently, and that they will select different environments when given the freedom to do so. All that sounds fine, but it also seems like a suite of proposals for which the crucial evidence would be standard policy intervention research, not supposed lessons from behavioral genetics.

The more direct intervention would be the use of genetics to inform the tailored education that children received. Asbury and Plomin make allusions to such possibilities. What makes precision medicine different from precision education is that it is far less clear how much value probabilistic polygenic information about individual students will have to offer above the information that can be obtained by measuring abilities and interests directly. Even if genomic data could (for example) reliably help identify learning challenges earlier in development and with higher fidelity, it is unclear that identifying challenges is really a key barrier to schools doing better. The present failure of schools to adequately tailor educational experiences to individual students seems more a failure of social priority than social science.

More broadly, outside of health, I am skeptical of the potential policy significance of genomic data being cast in terms of applications that use individuals' genomic information as a basis for intervention. This is usually what people leap to when imagining the potential import of genetic data: being able to use an individual's genetic information to act on that individual in some way. Instead, I think the main prospects are for policy-minded social science trying to understand better many of the things they are already studying. I am skeptical of the ultimate value of schools using children's polygenic scores to help them learn better, but I'm bullish on the possibility that polygenic scores can improve the statistical power of efforts to evaluate the effectiveness of school interventions and can help us better understand why interventions have different effects for different students.

What about eugenics?

In a nod to the famous ending of Michael Young's *The Rise of the Meritocracy*, Conley and Fletcher close *The Genome Factor* with an epilogue set a century from now. In vitro fertilization is used to generate a large number of fertilized embryos, and then polygenic scores give parents probabilistic information about these dozen or so embryonic siblings. Parents choose which one they wish to implant. Parents select the embryos with the most promising prospects for IQ; doing so confers such an advantage for social attainments that it becomes effectively normative; and population IQs rise to the point where IQs that are near-genius now are barely above average then.

As futurism goes, the technological requisites involved in this scenario are hardly fantastic; they may not even be far off. Political

barriers may well be imposed. In the United States, one can imagine strong objections from the political right at least as easily as from the left. Many speculate that some other societies—especially China and Japan—may be much less averse. The social barriers may be larger than what Conley and Fletcher's epilogue suggests. Parents' interest in optimizing prospects for something like intelligence may be more than offset if optimization comes with increasing risks of problems, and examples from artificial selection make this seem quite plausible. Plus, Conley and Fletcher's "designer babies" scenario would take several generations, during which time who knows what other technological changes might preempt this scenario.

Regardless, moral challenges posed by new reproductive technologies have been here for quite awhile and are only going to increase as knowledge advances. Perhaps no word is more evocative than "eugenics" for the historical harms tied to genetic thinking, bringing to mind involuntary sterilization, anti-miscegenation laws, rationalizations of genocide, and other atrocities.

If all we were talking about was whether aspiring parents can kickstart their kid's IQ by a few points using IVF, it would be one thing. The tougher moral questions will center on what role concerns about eugenics should play in the use and regulation of genomic technologies to prevent problems. Iceland currently has only one or two children per year born with Down syndrome, as a result of wide prenatal screening and nearly all mothers choosing to abort in response to a positive screen. The potential for such developments to reconfigure ideas about reproductive rights and social responsibilities to disabled children is obvious—here *Backdoor to Eugenics* (1990) seems especially prescient—even without the extra dystopian layer added by the private-sector domination and inequalities of the U.S. health care system.

I believe researchers working in social science genomics have a moral responsibility to familiarize themselves with the history of eugenics. Doing so, however, has taught me that many sociologists have serious misconceptions about this history. Specifically, presentism pervades how many people imagine the history of eugenics, inviting notions that eugenics was caused simply by ideas that good moral citizens today all recognize as bad. This is how the history of eugenics appears in *Social by Nature*, for example. A sobering book for understanding how wrong this view is—especially for reckoning with social science's legacy regarding the eugenics movement—is Thomas C. Leonard's *Illiberal Reformers*.

One sociologist with a significant role in *Illiberal Reformers* is E. A. Ross, the fifth president of the American Sociological Association. Ross coined the term "race suicide," described the Civil War as "lining the valleys of the South with the bones of half a million picked whites in order to improve the conditions of four million unpicked blacks" (p. 121), and was fired from Stanford after saying it would be better to sink incoming ships of Japanese migrants rather than let them in the United States. Ross was a key advocate—the year before ascending to ASA's presidency—for Wisconsin's involuntary sterilization law.

I do not expect any sociologist to be surprised that one of sociology's early leaders said things we now find appalling. But the convenient way to think about someone like Ross is that he offered a vision of sociology that was swept away by how properly enlightened sociologists think today. Instead, Ross was writing repellent things at the same time he was writing other things that would still read as orthodox sociology, not to mention orthodox leftism, today. Leonard's book also shows how eugenics arguments were used to support positions that nearly all sociologists today regard as Progressive Era triumphs. For example, one argument for establishing a minimum wage was that it would protect superior workers from being replaced by two inferior (read: immigrant) workers given a fraction of the pay. Contemporary social science is connected to a fraught and troubling history that is important for its researchers to recognize and reckon with, but the lessons of that history are not just relevant to sociologists whose work has something to do with genes.

Conclusion

Contemporary genomics is a gigantic global project that is propelled primarily by its potential health applications. The largest

genotyping capacity in the world is in China; the largest single data source is the UK biobank; the epicenter of important statistical advances has been Australia; the most extensively genotyped nation is Iceland. In the United States, the largest datasets are held by a private corporation that individuals have paid to have assay their saliva.

Social science genomics exists almost entirely in this slipstream. What would happen if every social scientist who is working with genomic data stopped work tomorrow? Available sample sizes will keep getting larger. Genotyping will keep getting cheaper and more extensive. Analytic tools will keep getting more sophisticated. Even papers on social science topics using genomic data would continue, as today many of the papers on social science topics are being written on the side by people whose bills are paid by health research.

Put simply, that social science genomics has a future is a sure thing. Far less certain is what part of that future will involve disciplinary sociology. Reader, I have no idea. So many broad uncertainties are afoot in disciplinary sociology: about the future of science-oriented research, of quantitative research, of grant-funded research, of policy-minded research, or research about individuals and why their life outcomes differ. Disciplinary sociology could well settle on an intellectual garden free from folks working with genomic data, with the only discussions of genomics being by those who do critical science studies about it.

Emblematic of that possibility might be *Social by Nature*'s argument that what is needed instead of social science genomics is a "gene-environment approach that turns the tables on gene-first science" (p. 219). Every indication Bliss offers as to what this might look like, however, involves only platitudes about structural causes and built environments. Tellingly, Bliss offers no ideas for how her new gene-environment approach would actually involve genes. Disciplinary sociology recognizes that overtly denying any influence of genes sounds retrograde, but it is unclear whether, except on its peripheries, the discipline will ever be able to bring itself to talk with serious specificity about genes.

It goes without saying that sociology will have no role in shaping emerging discussions of gene-environment interplay if the discipline cannot speak to both sides of that interplay. Social science genomics will continue forward either way, but it will be better if sociology is involved. Sit in enough interdisciplinary events and the blind spots of psychology and epidemiology become plain. *The Genome Factor* is the highest profile book about social science genomics to have either a sociologist or economist as a coauthor (much less both). Read it and a book like *G is For Genes* side by side, and you will see completely different ways of posing and thinking about questions. The coming power and promise of polygenic tools for interrogating social science questions is enormous, and their uses will be shaped by a vast, transdisciplinary conversation. For sociology to have any influence on how that conversation turns out, there need to be sociologists working fluently with genomic data.

Work Discussed in This Essay

Ashraf, Quamrul, and Oded Galor. 2013. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review* 103(1): 1–46.

Blackburn, Elizabeth, and Elissa Epel. 2017. *The Telomere Effect: A Revolutionary Approach to Living Younger, Healthier, Longer*. Hachette, UK: Grand Central Publishing.

Bliss, Catherine. 2018. *Social by Nature: The Promise and Peril of Sociogenomics*. Palo Alto, CA: Stanford University Press.

Comfort, Nathaniel. 2018. "CRISPR's Willing Executioners." *Nature* 553:278–80.

Conley, Dalton, and Jason Fletcher. 2017. *The Genome Factor: What the Social Genomics Revolution Reveals about Ourselves, Our History, and the Future*. Princeton, NJ: Princeton University Press.

Daw, Jonathan, Michael Shanahan, Kathleen Mullan Harris, Andrew Smolen, Brett Haberstick, and Jason D. Boardman. 2013. "Genetic Sensitivity to Peer Behaviors: '5HTTPLR,' Smoking, and Alcohol Consumption." *Journal of Health and Social Behavior* 54:92–108.

Dobbs, David. 2014. "The Fault in Our DNA." *New York Times*. July 10.

Duster, Troy. 1990. *Backdoor to Eugenics*. New York: Routledge.

Gould, Stephen Jay. 1981. *The Mismeasure of Man*. New York: Norton.

Graham, Loren. 2016. *Lysenko's Ghost: Epigenetics and Russia*. Cambridge, MA: Harvard University Press.

Lee, James J., et al. Forthcoming. "Gene Discovery and Polygenic Prediction from a 1.1-Million-Person GWAS of Educational Attainment." *Nature Genetics*.

Leonard, Thomas C. 2016. *Illiberal Reformers: Race, Eugenics, and American Economics in the Progressive Era*. Princeton, NJ: Princeton University Press.

Martin, Alicia R., et al. 2017. "Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations." *The American Journal of Human Genetics* 100(4):635–49.

Moore, David S. 2016. *The Developing Genome: An Introduction to Behavioral Epigenetics*. New York: Oxford University Press.

Nikpay, Majid, et al. 2015. "A Comprehensive 1000 Genomes-Based Genome-Wide Association Meta-Analysis of Coronary Artery Disease." *Nature Genetics* 47(10):1121–30.

Roberts, Dorothy. 2011. *Fatal Invention: How Science, Politics, and Big Business Re-create Race in the Twenty-First Century*. New York: The New Press.

Wade, Nicholas. 2014. *A Troublesome Inheritance: Genes, Race, and Human History*. New York: Penguin Group.