Genetic determinism versus genetic ignorance

BioSocieties,* Book Review Forum, pp. 2-9

Melinda C. Mills, University of Oxford and Nuffield College

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

Robert Plomin, *Blueprint: How DNA makes us who we are*, MIT Press, 2018

The field of behavioural genetics was founded in the late 19th Century, and embraced by the eugenics movement in the late 1880s, with the aim to 'improve' humanity via selective breeding. Countless people were sterilised and, in Nazi Germany, millions lost their lives. Modern behavioural genetics began in the mid-20th Century, focussing on heritability using family and twin-based studies. 'Heritability' estimates compared monozygotic (identical) and dizygotic (same genetic resemblance as siblings) twins to parse out the variation of a trait in a population attributed to genetic differences. Stating that something was 20% or 80% 'heritable' was the common take-home message from these studies. Heritability, as the term is used in this research, was and is, however, recurrently misunderstood. It is not about individuals, but rather differences within a single population. And low heritability does not necessarily equate to no genetic contribution, but rather absence of relevant genetic variance in a particular population. Think of height in a population struck by famine. Heritability of height may be low or even zero, but there is a still a genetic basis for height. Or, think of the number of vertebrae that you have. The reason is highly genetic, yet the number of vertebrae has low heritability since it lacks variance and is the same in everyone.

Genetics revolutionised at a staggering rate with the sequencing of the human genome in 2003, the plummeting of genotyping costs and the rise in computing power. After years in the wilderness, behavioural genetics researchers were suddenly able to utilise molecular genetic data in the form of single-nucleotide polymorphisms (SNPs). We are able to use SNPs for measurement since they are the single-base differences in DNA that mark genetic variation. The cheap measurement of SNPs enabled the birth of genome-wide association studies (GWAS) in 2005, which moved beyond a certain heritable percentage to isolate actual genes and in some cases, their biological function. A GWAS is a statistical search scouring the entire genome across each genetic variant to see if there is a statistical relationship between SNPs and a trait or a disease.

By 2019, almost 4000 GWA studies had examined thousands of traits, from height, Alzheimer's Disease and coffee consumption to the number of children born to an individual (Mills and Rahal 2019). In the mid-2010s, traits that were the classic purview of the social sciences joined these GWAS discoveries. This included educational attainment (Lee, Wedow et al.2018; Okbay, Beauchamp, et al. 2016; Rietveld et al. 2013), wellbeing (Okbay, Baselmans, et al. 2016) and a study I led on fertility (age at first birth, number of children) (Barban et al. 2016). Traits that had previously been explained by purely socially

* https://doi.org/10.1057/s41292-019-00156-6. Comment from Author - It states to use this doi to cite the book review editor only and not the authors of the reviews, but I will let you decide what is just for yourself.

deterministic theories were suddenly confronted with genetics and biology. A key discovery was that these traits were not explained by 'one gene' but rather a combination of many genes, each having a small effect. The results of GWA studies for complex traits are thus combined into what is called a polygenic score. This period also heralded the birth of 'sociogenomics', a term first used in relation to honey bees in 2005 (Robinson, Grozinger, and Whitfield 2005). Although Bliss's book is on sociogenomics, she leaves out this key fact as well as links between the study of human, animal behaviour and genetics.

It is important to offer this brief history here, because the books of Bliss and Plomin are about these transformations, with Bliss examining the science and scientists of sociogenomics and Plomin chronicling a lifetime in the field of behavioural genetics. They epitomise the faultlines and tensions. But what they have in common is that they both struggle to straddle and truly engage in the very discipline that they are critiquing or discarding. For Bliss it is ignorance of molecular and population genetics and for Plomin it is of sociology, family, educational and life course research. It admittedly took me some time to pick up Bliss' book because it was—rather uncomfortably—about me, a sociogenomics researcher, and about 'us', my colleagues in the field. An early review of the book in *Nature* did not help with its highly offensive title 'CRISPR's Willing Executioners' (pulled rapidly from the online version, but remains in the print version). For me, who lost family in WWII, my dear German collaborator and many others, cavalier Nazi references were not only unfair, but reckless. The sociologist Jeremy Freese, who has worked in the area of sociogenomics, has previously listed a litany of errors in the book such as confusing the 'null' hypothesis with 'no' hypothesis, but there are considerably more to note such as referring to a prominent genetics company as 23andWe (p. 48) or a female Canadian-Dutch sociogenomic researcher as an American. The apparent lack of attention to details such as these, attempts at sensationalism and other assumptions make the book a hard slog.

Beyond this, though, there is a serious problem of a deep misunderstanding of core terms in the field that Bliss is trying to critique. It was hard to grasp whether this was due to inadequacies in the research, or 'intentional ambiguity'. She presents numerous fuzzy definitions on key terms such as heritability or pleiotropy (i.e. when one gene influences two or more seemingly unrelated traits) or mixes up statistical methods (GREML) with the computer program that runs it (GCTA) (p. 78). Other claims are that sociogenomics "has risen to power largely under the radar of mainstream genomics" (p. 2) and that we developed our own methodology (p. 74). Publishing in the top genetics journals such as *Nature Genetics*together with geneticists seems hardly under the radar. Pre-analysis plans are on the Open Science Framework and articles are comparable with others. Sure, we might have a few more age-period-cohort models and econometric applications, but it is comparable.

It was hard to choose the most misinformed quote, but this is one example "…social genomics researchers allow the genomic software to tell them about the relationship between phenotypes. From there, they require no further socio-environmental analysis…. they even allow the software to determine how phenotypes play out along ethnic lines." (p. 78) I will unpack this in a moment but the strangest mistake is that she also states that the field focuses on gene–gene interaction, which is patently incorrect. What a shame that a book on sociogenomics entirely missed and actively downplayed the rich research of gene–environment interaction work done by Boardman, Fletcher, Belsky, Domingue and Conley to name just a few. What I read as Bliss' occasional disengagement, or a certain lack of substantive interest and understanding is summarised best as follows: "I sat in my seat at the back of the room, checking my e-mail as the last speaker of the morning shimmied offstage

and the next speaker walked on." (p. 140) Is it unfair to wonder if it might not have been better to listen to the content of the presentation?

As an expert in race and genetics, it is even more surprising that Bliss reiterates the mistake of equating race with population stratification (p. 90) or, as she states, "population stratification" is a "correction for managing race in study design" (p. 100). She also mischaracterises the focus on European ancestry populations as "ethical rationales" (p. 90) (see (Mills and Rahal 2019) for detailed empirical analysis). Many may recall the recent uproar over David Reich's (2018) book and subsequent heated and public discussion over race, ancestry and population stratification. After reading Reich's book myself I wondered how closely those who condemned him had actually attended to his book or attempted to understand this distinction. For example, on the first page he thanks his wife for researching the book and preparing the first drafts of the chapters together yet she is not a co-author. I would have expected that card-carrying social scientists would have drawn attention to this. None have done so, that I have seen. This is not to offend or detract from the content of Reich's book, but it harkens to work highlighting decades of female programmers in the footnotes of key genetic discoveries (Dung et al. 2019). Seeing this exciting precedent I asked my husband to do the research and help me write my next book—he declined.

Let us put the record straight. Race is not a biological or genetic category but rather a social construct with a strong social, cultural and political meaning. (Bliss may also be surprised to learn that it is a term generally only used in the US, with ethnicity, for instance, favoured elsewhere.) Geneticists do not study race but rather population structure, which refers to the patterns that are found in genetic data that identify an individual's ancestry (not race) in relation to genetic admixture (interbreeding of populations which introduces new genetic lineages). The most common way to measure and detect population structure is by employing principal component analysis (PCA), a statistical technique that emphasises variation and brings out strong patterns underlying the data.

Sociologists may relate to the use of PCA since it is similar to Pierre Bourdieu's model of the social as spatial and relational, captured by the term 'field' (Bourdieu 1984). PCA is a dimensionality reduction technique. Just as Bourdieu's agents and social positions are located along multiple dimensions, so do genes mirror geography (Novembre et al. 2008), with rich descriptions identifying how human migrations or historical events such as the Mongol empire or European Colonialism have shaped human populations.(Hellenthal et al. 2014) Evidence places the origins of Homo sapiens in what is now Sub-Saharan Africa (Namibia and Angola) since patterns of DNA sequence variation are the largest there (i.e. the greatest genetic diversity is found in sub-Saharan Africa). In genetics, the greatest genetic variation is assumed to be the oldest. Even two people from adjacent villages in indigenous groups in southern Africa have more variation than would be shown in a comparison between two people of European and Asian ancestry. (Schuster et al. 2010). So to summarise, race is a socially constructed non-biological term and population stratification is about the dispersal and migration of populations, linked with geography. Bliss states that controlling for population stratification is controlling for race. But race in the American sense she refers to has already been removed before we control for population stratification.

On to the next book: Robert Plomin has had an incredible career within psychology and is one of the founders of behavioural genetics. Contrary to Bliss, he provides an accessible and clear definition of key terms such as polygenicity and pleiotropy. It was fascinating to read a first-hand account of the pioneering studies he led such as the Colorado Adoption Project and

the Twins Early Development Study (TEDS), which truly advanced the field. In the first part of the book he targets many of the misnomers and discussions in the past and present such as heritability and the *p* value discussion. He walks us through milestones such as the 'aha' moment when multivariate regression models could be integrated into twin models. All of these stories, disappointments and victories provide a deeper understanding of the incredible span of his academic career and the many challenges he faced. However, although I have deep respect for his work, I could not help but thinking when I read the book that it was almost *The Strange Case of Dr. Jekyll and Mr. Hyd*e, where the good doctor joins the dark side. Some of the statements are so surprising that he even contradicts himself and his own previous work. Many of the statements are bold and outrageous and although he dials it back in the endnotes, the core message remains.

There are two problems: the first is that Plomin claims that genetics trumps everything *and*that we can use these new polygenic scores for precision education policy. He argues that genes are king, the 'genomic genie is out of the bottle' and that all non-genetic effects (families, schools, life experiences) are non-systematic and random and by extension do not have an effect. For a man who coined the term 'non-shared environment', he has an outlandish misunderstanding of how it works. Just as Bliss fails to engage with genetics, Plomin falls into the same trap with sociological, educational and life course research. Plomin's main argument is that the non-shared environment is not only unsystematic and due to chance, but also unstable, random and inconsistent over time or what he refers to as 'random noise' (p. 80). After searching for 30 years he could not find it, so it must not be there. Here we read several sections such as 'Parents matter, but don't make a difference' to 'Schools matter, but….' and 'Life experiences matter, but…'. Why is this problematic?

First, if you want to understand the importance of the social structural environment, do not hire a psychologist working in the area of genetics. The methods focus on variance and not means, so by virtue of this approach, inherently ignore group differences. The data used are also narrow and rarely capture environmental variation. Research in sociology, education and the life course has a rich history of demonstrating the systematic and long-term cumulative advantages and disadvantages of family, education and life course events. Second, even if we guaranteed that families, the educational system and our life experiences were 'random noise', why cannot it still have an effect? Third, Plomin often returns to the empirically thorny ground of gene–environment correlation, arguing that we select, modify or even create our own environments. The main tenet of the mantra "parents matter, but don't make a difference" is that "children make their own environment, regardless of parents." (p. 83). And further: "children who want to do something like play sports or a musical instrument will badger their parents to make it happen." (p. 83) Or later: "opportunities are taken, not given" and it would be a "mistake to see it as inequality" (p. 96) because in the end it is all based on genetics.

No. This just is not true. Children from disadvantaged backgrounds cannot "badger" their parents for sports or music lessons until they get them nor can we blame them for not "taking" opportunities that were never there for them in the first place. Countless studies have shown socioeconomic disparities in extracurricular activities from young ages right up until high school where students from wealthier backgrounds are able to enhance their activities, further widening the admissions competition and higher education divide.

What about schools, which Plomin characterises as the random chaos where children spend over a decade of their lives? This and Chapter 9 on the inequality of opportunity and

meritocracy is a grim read for a sociologist. Plomin discusses meritocracy as selection based on capability and competence. He foresees a brave new world where people are judged by their 'innate' genetic abilities. Digging an even bigger hole with intergenerational mobility, he states that "parent–offspring resemblance is an index of heritability, and heritability is an index of equal opportunity" and also that "….parent–offspring resemblance for education and occupation indicates social mobility rather than social inertia" (p. 95). The problem with this is that genetics has shown us that there is considerable variation amongst siblings in the same family (Visscher et al. 2008). Siblings each randomly inherit half of their parent's DNA, but these are different halves. There is segregation of genes within families, so if a trait such as height has a heritability of 0.8 and there is a standard deviation of 7 cm in a population, the standard deviation of children around the mean height of parents is 5.4 cm ($= \sqrt{[7^2(1 - ½*0.8)]}$). Tall parents have on average tall children but there is quite some variation around the mean (Visscher et al. 2008).

Surprisingly, for someone who has lived in the UK since the early 1990s, Plomin seems to have a complete lack of contextual knowledge and states that "in the UK there is little selection because most parents send their children to a local school" (p. 97) and by virtue of this "children at different schools receive equally good education". In the UK, the school you go to is determined by what is called a 'catchment area' or in other words by where you live, which is highly socioeconomically stratified. He does note the difference in school achievement between state-funded (what he terms 'non-selective') schools and what he calls 'selective schools' (private schools, in the UK sometimes, rather confusingly, known as public schools). The reason that grades are an entire grade point higher in private schools, he argues, is purely based on their meritocratic "competitive selection" of students and for this reason "students in selective and non-selective schools differ in their DNA" (p. 98). He rather incredibly fails to mention the prohibitive costs of these schools that only the wealthy can afford. Admission to some of the top private boys' schools is around £42,000 ($55,000 USD) per annum for fees alone (and if the boys 'badger' their parents for music and sports a few more thousand on top of that). That would take some momentous badgering for a kid living in a publically supported council house or a family trying to make ends meet, in Plomin's world where "opportunities are taken, not given".

In his meritocracy argument, Plomin also focuses almost entirely on test scores. Indeed diverse schools teach similar skills, but what Plomin misses is that meritocracy is also the acculturation, status, peer and social network effects of these schools. The power of schools is far beyond test scores. A recent article demonstrated that UK elite schools are the 'nurseries' of the old-boys network of the powerful British elite (Reeves et al. 2017). Sociologists have shown that free schooling reduces inequalities in educational attainment (Breen 2010) but also how parents in more advantaged class positions actively help children to maintain their competitive edge in the educational system (Bukodi and Goldthorpe 2018). There is a vast amount of evidence of the persistence of inequality of educational opportunity in the UK (Boliver 2013). It is not only parental background or school programmes that enable children to fully develop their educational potential, but a broader range of economic and social policies that matter.

The second problem is when Plomin argues for the use of polygenic scores of educational attainment as social policy instruments to extend the idea of "precision medicine" to "precision education". At the end of the book, he describes polygenic scores by looking at the distributions of his own scores. He de-emphasises the shotgun scatterplot of these polygenic scores to focus on the tails of the distribution. There he finds that he is in the 94th percentile

of educational attainment. Congratulations. What he does not mention is that a lot of people end up on the left hand of the distribution, including Yaniv Erlich, who—rather ironically—developed the code Plomin used to show he was in the 94th percentile and science writer Carl Zimmer. (For an exceptionally entertaining read on this topic see: Zimmer 2018.)

There are multiple impediments to applying Plomin's "precision education" score, and I describe here only the obvious ones that Plomin fails to address. First, around 90% of polygenic scores are derived from European ancestry populations and 72% are recruited from just three countries (UK, US, Iceland) (Mills and Rahal 2019). Polygenic scores derived from European ancestry groups cannot be applied to non-European ancestry groups (Martin et al. 2017). Or as we noted previously, these genetic discoveries that the polygenic scores are based on miss 76% of the world's population (Mills and Rahal 2019). Using these scores would be for the healthy, European ancestry groups and serves only to exacerbate inequalities and must cross lines of basic human rights. Second, due to the ubiquity of pleiotropy, the causal function of the genes in the polygenic scores is mostly a black box. We simply do not know how fragile they are or when they stop working. Not only are they at risk of being ineffective, but could even have unintended or harmful consequences. Third, the prediction is far from "precision" (see Kaufman 2019 on this). If someone scores at the 75th percentile of the polygenic score for educational attainment, we can predict that they will score somewhere between the 2nd and 98th percentile chance of academic achievement. If someone told you just before you were going to board a plane that it had between a 2 and 98 confidence interval of going down, would you get on? I don't think so. Even if we took the individuals at the high end of the tail (90-100th percentile), they still on average would reach the 66th percentage and score between the 15th and 100th percentile. Still interested in getting on that plane?

Building on Troy Duster's claim that sociogenomics is a "backdoor to eugenics", Catherine Bliss concludes that researching the genetic basis of social behaviour such as educational attainment and sociogenomics is a "frontdoor to eugenics". Recall that eugenics is not about applying a genetic score for policy, rather it means altering groups by (dis)encouraging reproduction of people or populations with certain traits. Although Plomin's genetically engineered "precision education" is absurd, neither he nor any others in the sociogenomics field are arguing to link "desirable traits" to reproduction. Beyond the obvious and completely repugnant nature of this thought, technically it is simply ludicrous and would suffer the same problems as Plomin's precision education I listed above.

Plomin argues that this persistent nature-nurture debate in the social sciences "was driven by misplaced fears about biological determinism, eugenics and racism" (p. 53). I can only think that Plomin's overzealous claims are driven by someone who spent 30 years under attack by sociologists armed with the kinds of misguided arguments that Bliss proposes, and at the end of his career finally saw a possibility of molecular genetic predictors. Given Plomin's high scientific stature, Bliss does get at something we need to be concerned about, however. Governments often need a quick and inexpensive silver bullet based on 'science' to solve health and educational disparities. Beyond obvious diversity and methodological issues, a focus on these genetic predictors alone distracts us from the real problem and interventions at the family, school, neighbourhood and national level.

Acknowledgements The author would like to thank Des Fitzgerald, Cathy Waldby, Charles Rahal, David Brazel and John Goldthorpe for comments on this review.

References

Barban, Nicola, et al. Melinda C. Mills. 2016. Genome-Wide Analysis Identifies 12 Loci Influencing Human Reproductive Behavior. *Nature Genetics* 48 (12): 1–7.

Bukodi, Ersebet and John Goldthrope. 2018. *Social Mobility and Education in Britain: Research, Politics and Policy.* Cambridge: Cambridge University Press.

Boliver, Vikki. 2013. How Fair Is Access to More Prestigious UK Universities? *The British Journal of Sociology* 64 (2): 344–364.

Bourdieu, P. 1984. *Distinction: A Social Critique on the Judgement of Taste*. London: Routledge.

Breen, R. 2010. Educational Expansion and Social Mobility in the 20th Century. *Social Forces*89 (2): 365–388.

Dung, Samantha Kristin, Andrea López, Ezequiel Lopez Barragan, Rochelle-Jan Reyes, Ricky Thu, Edgar Castellanos, Francisca Catalan, Emilia Huerta-Sánchez, and Rori V. Rohlfs. 2019. Illuminating Women's Hidden Contribution to Historical Theoretical Population Genetics. *Genetics* 211 (2): 363–366.

Hellenthal, G., G. B. J. Busby, G. Band, J. F. Wilson, C. Capelli, D. Falush, and S. Myers. 2014. A Genetic Atlas of Human Admixture History. *Science* 343 (6172): 747–751.

Kaufman, Scott Barry. 2019. There is No Nature-Nurture War, Scientific American, Jan. 18, 2019.

Lee, J. J., R. Wedow, and Okbay. 2018. Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals. *Nature Genetics*.

Martin, Alicia R., Christopher R. Gignoux, Raymond K. Walters, Genevieve L. Wojcik, Benjamin M. Neale, Simon Gravel, Mark J. Daly, Carlos D. Bustamante, and Eimear E. Kenny. 2017. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *The American Journal of Human Genetics* 100 (4): 635–49.

Mills, M. C. and C. Rahal. 2019. A Scientometric Review of Genome-Wide Association Studies. *Communications Biology* 2 (9).

Novembre, John, Toby Johnson, Katarzyna Bryc, Zoltán Kutalik, Adam R. Boyko, Adam Auton, Amit Indap, Karen S. King, Sven Bergmann, Matthew R. Nelson, Matthew Stephens, and Carlos D. Bustamante. 2008. Genes Mirror Geography within Europe. *Nature* 456 (7218): 98–101.

Okbay, Aysu, et al. David Cesarini. 2016. Genetic Variants Associated with Subjective Well-Being, Depressive Symptoms and Neuroticism Identified through Genome-Wide Analyses. *Nature Genetics* 48 (6): 624–33.

Okbay, Aysu, et al. Daniel J. Benjamin. 2016. Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment. *Nature* (6139): 1467–1471.

Reeves, Aaron, Sam Friedman, Charles Rahal, and Magne Flemmen. 2017. The Decline and Persistence of the Old Boy: Private Schools and Elite Recruitment 1897 to 2016. *American Sociological Review* 82 (6):1139–66.

Rietveld, C. A., et al. P. D. Koellinger. 2013. GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science* 340 (6139): 1467–71.

Robinson, Gene E., Christina M. Grozinger, and Charles W. Whitfield. 2005. Sociogenomics: Social Life in Molecular Terms. *Nature Reviews Genetics* 6 (4): 257–270.

Schuster, Stephan C., et al. Vanessa M. Hayes. 2010. Complete Khoisan and Bantu Genomes from Southern Africa. *Nature* 463 (7283): 943–947.

Visscher, Peter M., W.G. Hill and Naomi R. Wray. 2008. Heritability in the Genomics Era—Conceptions and Misconceptions. *Nature Review Genetics* 9 (4): 255–266.

Zimmer, Carl. 2018. Genetic Intelligence Tests Are Next to Worthless. *The Atlantic*.