Confounding Fuels Hereditarian Fallacies

John W. Benning[1]*, Jedidiah Carlson[2,3]*, Ruth G. Shaw[4], Arbel Harpak[2,3]
[1]Department of Botany, University of Wyoming, Laramie, WY, USA
[2]Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA
[3]Department of Population Health, University of Texas at Austin, Austin, TX, USA
[4]Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN, USA

*these authors contributed equally
Correspondence should be addressed to: jwbenning@gmail.com, jedidiah.carlson@utexas.edu, shawx016@umn.edu, and arbelharpak@utexas.edu

Abstract

Scientific literature has seen a resurgence of interest in genetic influences on socioeconomic outcomes. Such investigations are often limited by confounding between signals of genetic and non-genetic influences. An illustrative example is Clark (2023), which considers the similarity in socioeconomic status between relatives, drawing on genealogical records spanning four centuries in England. Based on the fit of a quantitative genetics model, it suggests that social status is largely determined by one's DNA; and that, for that reason, contemporary English people "remain correlated in outcomes with their lineage relatives in exactly the same way as in preindustrial England." These conclusions are based on a conflation of genetic and non-genetic transmission (e.g., of wealth) within families. We demonstrate that additional errors and statistical artifacts influenced inferences in Clark (2023). In reality, Clark (2023) provides no information about the relative contribution of genetic and non-genetic factors to social status. We discuss how lessons learned from the failure to account for confounding generalize to contemporary studies that claim to establish genetic underpinnings to social outcomes.

Introduction

People vary remarkably in behavior and social outcomes, and this variation sparks curiosity about its causes. Scholars have debated the extent to which this variation among individuals arises due to underlying genetic variation for the past ~150 years. Variation in socioeconomic outcomes (e.g., educational attainment, wealth, occupation) has been a particularly intense focus of this debate.

Teasing apart genetic and non-genetic contributions to social outcomes is notoriously fraught. Galton (1) found strong resemblance of offspring to parents in measures of social status and from this inferred that genetics must be the primary driver, a school of thought described broadly as "hereditarianism" (see 2). As is now well-appreciated, Galton's inference ignored the fact that relatives share not only genes, but also wealth, place of residence, knowledge, religion, culture, and more. For many of these non-genetic factors, transmission within families can parallel genetic transmission (**Fig. 1**) (3–20). When those attributes that are non-genetically transmitted contribute to variation among people, their relative contribution is indiscernible from that of genetics. A long history of scholarship has highlighted this confounding and how it impedes inference of the causes of phenotypic variation, especially when genetic data are unavailable (21–28).

Nevertheless, there has in fact been a resurgence in hereditarian arguments, as the advent of Genome-Wide Association Studies (GWAS) brought about new opportunities to attribute socioeconomic variation to genetic variation (29). The failure to reckon with confounding may also be amplified by trends in publishing culture that increasingly incentivize sensationalization, citability, and media engagement.

To provide an illustration of this line of work and its limitations, we focus on a recent publication (30) as a case study. We identify the failure to account for confounding and two other core flaws in this study. We then discuss the general relevance of these flaws to studies linking genetic variation to socioeconomic outcomes.

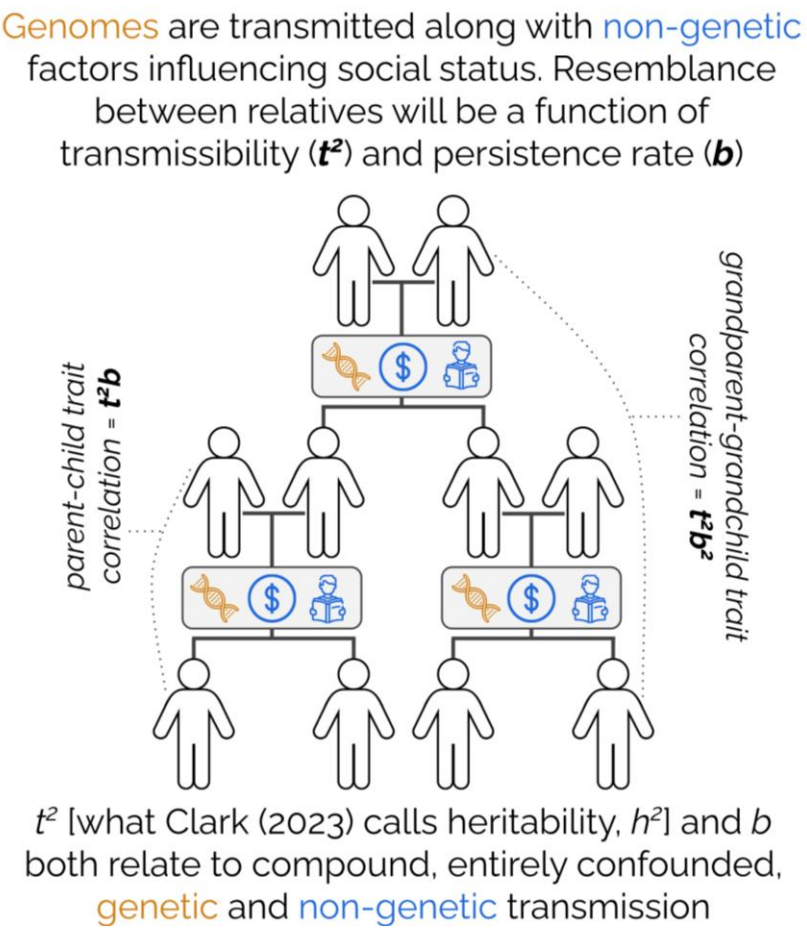

Figure 1. Non-genetic transmission can parallel genetic transmission, and their respective effects are confounded in observational data. Clark (2023b) uses a model where a trait value is the sum of an inherited component from parents and random noise. Under this model, the expected resemblance between relatives depends on transmissibility (t^2, the portion of trait variation attributable to the transmitted component) and a rate of decay across generations (the "persistence rate," *b,* which will increase as assortative mating increases). Ignoring the confounding of genetic and non-genetic transmission in the data, Clark (2023b) misassigns all transmission as genetic heritability and all assortative mating to be on a latent "social genotype".

Case study: Clark (2023)

(30) analyzed familial correlations in a dataset of socioeconomic measures (e.g., occupational status, house value, literacy) from a selection of English relatives spanning the 18th to 21st centuries. From analyses fitting these observed correlations to a quantitative genetic model of trait inheritance [(31, 32); **Supplementary Note 1**], (30) infers that social status persists intergenerationally because parents mate assortatively on a status-determining genotype (or "social genotype" as used by the author in previous work (33)). (30) then argues that because

mates share the genes underlying social status to such a high degree, the persistence of social status within families—and persistence of differences in status among families—have been largely unaffected by changes in social policy in the last four centuries. In a recent commentary about this work (34), the author presents the results of (30) as providing strong support for a hereditarian interpretation. In doing so, he appeals to the metaphor of a "genetic lottery" underlying social outcomes, a conceptualization increasingly in vogue in social and behavioral genetics (35) [critiqued in (36–39)], in order to argue that social status is determined by genetics.

Here, we discuss three core flaws regarding these claims (see our discussion of other misinterpretations, errors, and incongruencies in (30) in **Supplementary Notes 2-7**; **Tables S1-S3**; **Figs. S1-S13**). First, and most importantly, we explain that (30) fails to confront the confounding of genetic and non-genetic transmission (**Fig. 1**). We show an example of strong confounding between these two modes of inheritance in the data (**Fig. 2**). Second, we show that the estimated decay in correlations across genealogical relationships is partly explained by statistical artifacts (**Fig. 3**). Third, we demonstrate that familial correlations varied substantially over the time period examined, generally decreasing (**Fig. 4**). This finding stands in contrast to the paper's inference that familial correlations have been stagnant since the 17th century, and subsequent claim that there has been no change in social mobility. In summary, the data and analyses in (30) do not establish the contribution of genetics to social status.

Confounding between genetic and non-genetic transmission. Inferences in (30) are based on a linear regression model derived from quantitative-genetic theory developed by R.A. Fisher (31, 32) (**Supplementary Note 1**) and the model

$$P = G + E \qquad \text{Eq. 1}$$

where an individual's phenotype, *P*, is the sum of separable genotypic (*G*) and environmental (*E*) influences on it. Since genotypes are transmitted in families, genetic parameters can be inferred from correlations between relatives, as long as environmental influences can validly be assumed independent and random with respect to genotypes (by, for example, experimentally randomizing genotypes over environments). Fisher (1918) formally showed that under this model, the correlation in a trait between relatives is expected to be the product of the trait's heritability (h^2) times a compound parameter b that depends on the genealogical relationship between the relatives and the extent of assortative mating. (h^2 is the fraction of phenotypic variance due to additive genetic variance, nowadays referred to as "narrow-sense" heritability).

Crucially, to interpret the model parameters h^2 and b as genetic, Fisher's model makes the assumption that the phenotypic resemblance between relatives is purely due to a genetically heritable component. In particular, it assumes there are no non-genetic (material, environmental, or cultural) influences on a trait that are systematically shared or transmitted between relatives. In the application by (30), for example, it is assumed that similarity in house value (one of the measures of social status analyzed) is strictly due to shared genes, and does not arise from similarity in parental wealth between those relatives, or from the inheritance of wealth or property, or from having learned from one's relatives how to invest.

In reality, however, non-genetic transmission is ubiquitous for social and behavioral traits. Mechanisms of non-genetic transmission include "ecological inheritance," i.e., the trait value of an offspring is directly influenced by the environmental conditions created by their parents (e.g., familial wealth influencing educational opportunities) (8, 40), and the diffusion of information directly to one's relatives (e.g., literate parents teaching their children how to read) (5). When genotypes cannot be randomized over environments, true genetic effects are inseparable from other factors underlying phenotypic resemblance between relatives. In this regard, it is informative to consider the logic of genome-wide association studies (GWAS) which detect associations between traits and genetic variants (41, 42). The gold standard method to adjust for confounding in GWAS involves regressing out phenotypic resemblance that tracks genomewide genetic relatedness (14, 17, 43). In contrast, (30) assumes *a priori* that this signal of phenotypic similarity correlating with relatedness—well known by geneticists to be perfectly confounded—reflects only genetic effects.

In point of fact, there are signals of strong confounding between genetic and non-genetic contributions to familial resemblance in the data used in (30). The paper acknowledges the inheritance of material wealth from one's parents as an obvious example of non-genetic transmission—but only in treating wealth itself as the focal status measure. For other measures studied, the effect of familial wealth on social status writ large is ignored. Familial wealth can influence a wide range of conditions that affect offspring (e.g., healthcare, place of residence, access to tutors, social circles, etc.) (44–48). Thus, we were not surprised to discover that all seven status measures analyzed in (30) are substantially correlated with an individual's father's wealth (Pearson *r* ranging from 0.19 - 0.66; mean *r* = 0.36; all $P < 2 \times 10^{-16}$; **Table S2**; **Fig. 2a**). Closer relatives tend to have more similar paternal wealth, and the similarity in paternal wealth between relatives predicts their similarity in occupational status extremely well (Pearson *r* = 0.91;

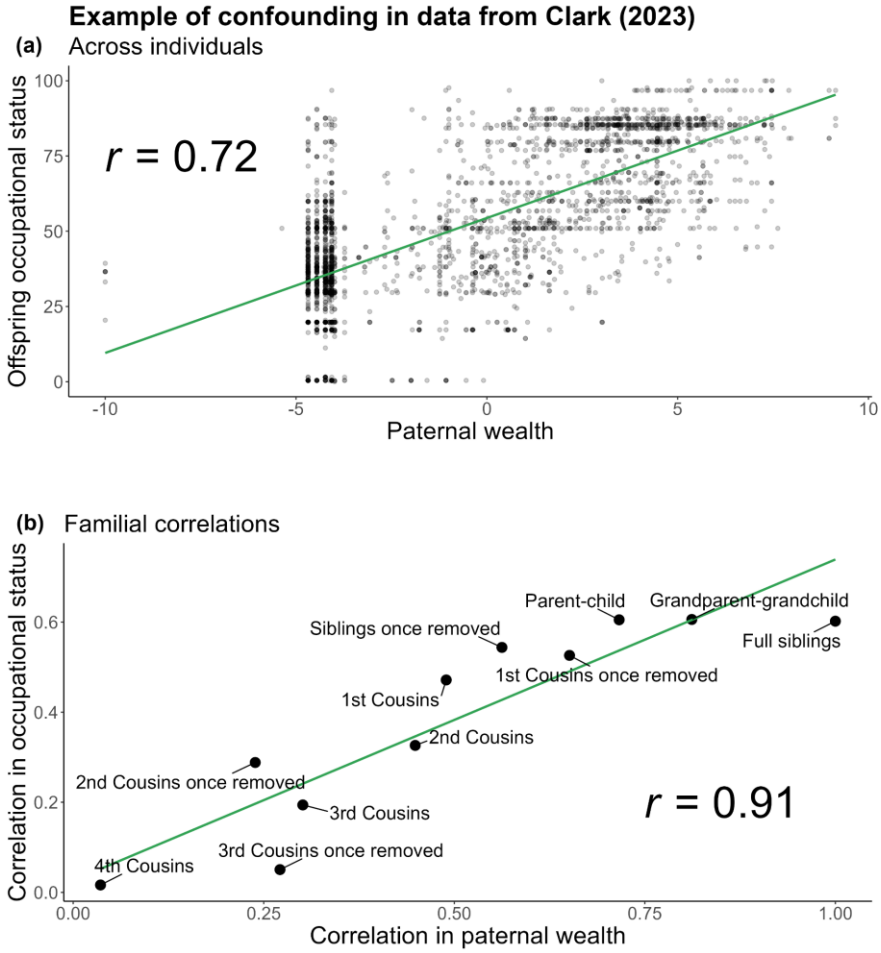

Figure 2. Confounding between sharing of genes and sharing of familial wealth in data from (30)**. (a)** An individual's occupational status is strongly correlated with their father's wealth. This relationship suggests at least one potential source of confounding between genetic and non-genetic transmission. **(b)** Correlations between relatives in occupational status are highly correlated with those relatives' correlation in paternal wealth. Plots show data for individuals born 1780-1859 for which paternal wealth data was available. N = 2,886 individuals in panel (a) and N = 13,030 pairs in panel (b). ((30) estimated wealth from probate records. The log of estimated wealth was mean-centered with respect to 5-year bin means. Individuals not probated due to insufficient wealth were assigned a value of half the minimum probate requirement for the time period.)

Fig. 2b). These analyses demonstrate the confounding in these data between transmission of genes and the effects of parental wealth on familial similarity in social status.

Numerous other non-genetic factors, apart from wealth, may contribute to familial correlations (49, 50). This confounding prevents genetic and non-genetic sources of familial resemblance from being disentangled. (30) presents two post hoc analyses as an attempt to rule out non-genetic

contributors to familial resemblance in status. In **Supplementary Note 4**, we detail why these analyses are actually uninformative as to the strength of non-genetic effects on resemblance in social status between relatives.

The confounding of genetic and non-genetic transmission in these data means that the interpretation of the model parameters offered in (30) as pointing to identifiable genetic contributions is misleading (**Supplementary Note 1**). In the presence of such confounding, the interpretation of G and E in **Eq. 1** as heritable genetic and random non-genetic effects on a phenotype, respectively, no longer holds. Instead, they can at best be interpreted as a transmissible component and a random, non-transmissible component. Consequently, the parameter interpreted in (30) as narrow-sense heritability, h^2, is in fact an estimate of the "total transmissibility" of a trait, t^2, the proportion of variance attributable to an unknown compound of transmissible effects including genes, culture, wealth, environment, etc. (10, 12). The second key parameter, m, which (30) interprets as the "spousal correlation in the underlying genetics," does not represent a genetic correlation between mates. It is instead the spousal correlation in the transmissible component of the trait. m is derived from the "intergenerational persistence rate," $b = \frac{1+m}{2}$, estimated from the regression model. The expected correlation for a given kinship pair is equal to $t^2 b^n$, where n denotes genealogical distance (**Fig. 1**). [Note that the parameterization of b for father-son and grandparent-grandchild relationships also depends on the degree of assortative mating with respect to the focal trait itself; see **Supplementary Note 1**]. This conflation of genetic and non-genetic transmission helps explain why the model parameters in (30) that are claimed to represent quantitative genetic parameters, h^2 and m, are much higher than estimates of these parameters in other studies that attempt to account for confounding (e.g., 19, 51; Supplementary Note 1, 52).

Statistical artifacts distorted estimates of familial correlations. In our reanalysis of (30)'s data, we found major statistical artifacts that influenced the main conclusions of the study. As one example, the analysis treats all pairs of relatives as independent observations. However, in (30)'s correlation analyses, many individuals are represented in multiple data points contributing to the estimate of a given correlation coefficient (**Fig. S7; Supplementary Note 6**). For example, the (1780-1859) occupational status correlation for fourth cousins is calculated from 17,382 pairs, derived from only 1,878 unique individuals from just 31 surname lineages. This is commonly known as pseudoreplication, with a known effect of driving underestimates of statistical

uncertainty (53). Here, however, the pseudoreplication is non-uniform: as genealogical distance increases, individuals are increasingly re-counted in more relative pairs, but these individuals are from a diminishingly smaller selection of lineages (**Fig. S7**, **Fig. S8; Fig. S9; Supplementary Note 6**). We note that individuals comprising more distant relationship types are also wealthier and have higher status (**Fig. S9; Fig. S10; Supplementary Note 6**). This trend may be partially due to sampling biases or temporal change in the population distribution of occupational status that are unaccounted for (**Fig. S10**). Because of the non-uniform pseudoreplication, the point estimates of familial correlations in (30) may even be biased. Indeed, when we avoided pseudoreplication by randomly sampling pairs of relatives (one pair from each surname lineage), point estimates of sample correlations do not even decrease monotonically with genealogical distance; in addition, their noisiness prevents making confident assertions about the persistence of these correlations with genealogical distance (**Fig. 3; Supplementary Note 6**). Together, the core flaws of model misspecification and statistical artifacts call into question inferences drawn in (30) based on familial correlations in status.

A reappraisal of the 'persistence rate' as a measure of social mobility in Clark (2023). Claims about the insensitivity of familial correlations to social interventions in (30) rest on the paper's finding that the parameter b ("the persistence rate of the correlation as we move one step down the family tree, or one step across between full siblings") is similar across status measures and across time. For example, if the correlation in occupational status between first cousins is 0.8 that of uncles and nephews, which is 0.8 that of full siblings and so on, then $b = 0.8$ (**Fig. 1**; also see Table 2 in (30)). As discussed in **Supplementary Note 1** and above, in (30)'s implementation of (31)'s generative model, b is a deterministic function of the assortative mating parameter m, the correlation between spouses in the transmitted component of the trait.

(30) argues that b has been stable across time, status measures, and families—and that this stability is due to strong assortative mating on a genetic factor for "social ability", estimated as a genetic correlation of $m = 0.57$ between mates. However, once one acknowledges that both genetic and non-genetic factors are transmitted within families (**Supplementary Note 1**; **Figs. 1, 2**), it follows that m tells us nothing about genetic versus non-genetic contributions to assortment, and b tells us nothing about the cause of within-family persistence of social status.

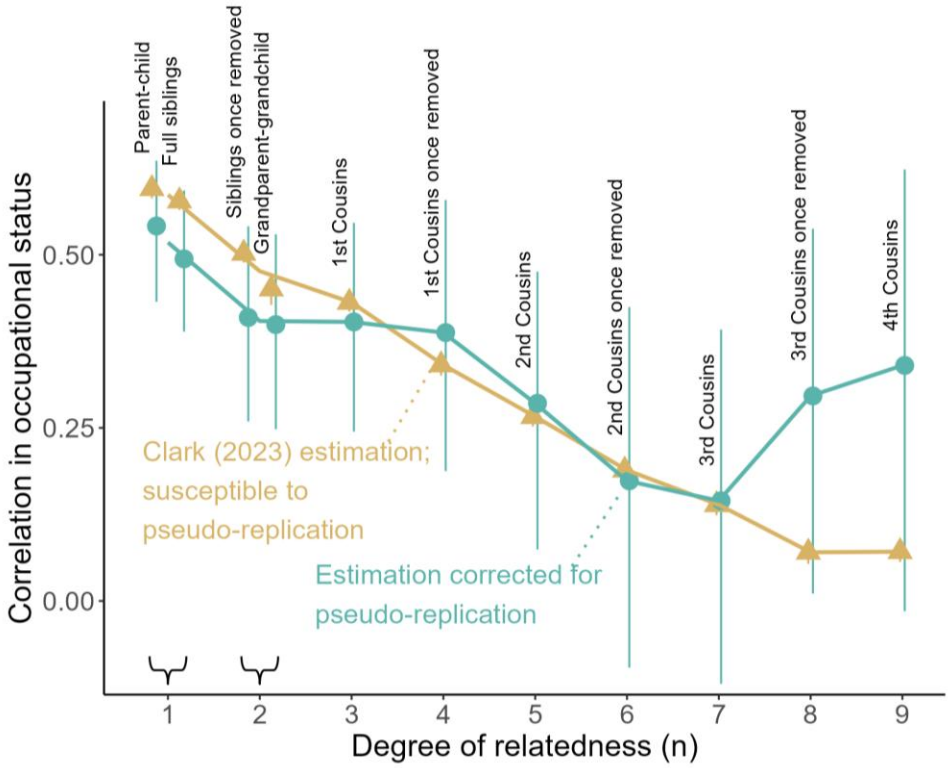

Figure 3. **Pseudoreplication distorted estimates of familial correlations.** Familial correlations (95% CI) in occupational status (1780-1859) using the approach employed by (30) (in gold) involved pervasive, non-uniform pseudoreplication. In contrast, in teal we show conservative estimates using only a single relative pair per surname [means and 95% CI over 1000 bootstrap samples are plotted for each familial correlation], which are therefore not susceptible to pseudoreplication. Distant cousins show dramatically higher correlations after adjusting for pseudoreplication.

Regardless of the cause, (30) claims to present evidence that

> "*The vast social changes in England since the Industrial Revolution, including mass public schooling, have not increased, in any way, underlying rates of social mobility*".

If supported by the data, this would be a striking finding. This claim rests on the observation that *b* is similar between two time periods for two status measures. However, between the two time periods analyzed, 16/22 familial correlations decrease (on average, decreasing 31%) [(30) Table 2]. How could the estimate of b lead to such contrasting conclusions? (30) offers neither justification for b as a measure that bears on social mobility, nor explanation for ignoring established measures of mobility and discrepancy with other literature (e.g., 54). In the data analyzed by (30), parent-child correlations (a common measure of social mobility) in occupational

status, higher education, and literacy generally decrease over time (**Fig 4; Supplementary Note 7**).

Some readers have already taken arguments in (30) as compelling evidence that social status is largely caused by genetic factors (55–57). Yet the assumptions and interpretations in (30) ignore a century of quantitative-genetic theory, previous empirical evidence for confounding, and the fallacies that arise when confounding is ignored (13, 17, 21, 22, 26, 37, 38, 58–62), as well as patterns in the paper's own data that conflict with the interpretations presented. (30) does not merely overstate the findings—the model parameters are misconstrued and the pervasive confounding of genetic and non-genetic transmission in these data is not interrogated.

Relevance to contemporary genomic studies

Many of the claims made in Clark (30) seem to rely on an implicit assumption that transmission in families is solely genetic. The analysis, based only on observational phenotypic data, is similar in spirit to ones that could have been carried out by Francis Galton a century and a half ago. Are the inferential flaws described above relevant to contemporary studies that use large genomic datasets and employ state-of-the-art statistical methods to adjust for confounding? We posit that these concerns are still broadly relevant, because confounding is still poorly understood and often underplayed in the literature.

Residual confounding in genomic studies remains poorly understood. Human geneticists have long appreciated that there are myriad ways by which a genetic variant may be associated with a trait or outcome (26, 63, 64). Accordingly, in analyzing data (e.g. in GWAS), they strive to adjust for confounding with various methods. However, residual confounding (i.e. confounding that persists even after application of these statistical methods) may still bias estimates. In 2019, we and other researchers discovered that genetic effect estimates in the largest GWASs for height—the most extensively studied polygenic human trait—were biased due to residual confounding. It became clear that the bias for each individual genetic variant was slight, but it was systematic across variants. Consequently, when researchers summed over signals from many genetic variants, they also summed over systematic biases. This led to erroneous conclusions in numerous studies (as detailed in 17, 60, 61). Further research has demonstrated that residual confounding may affect many GWASs, in particular for social outcomes and traits that are mediated by social context (26, 62, 64–67).

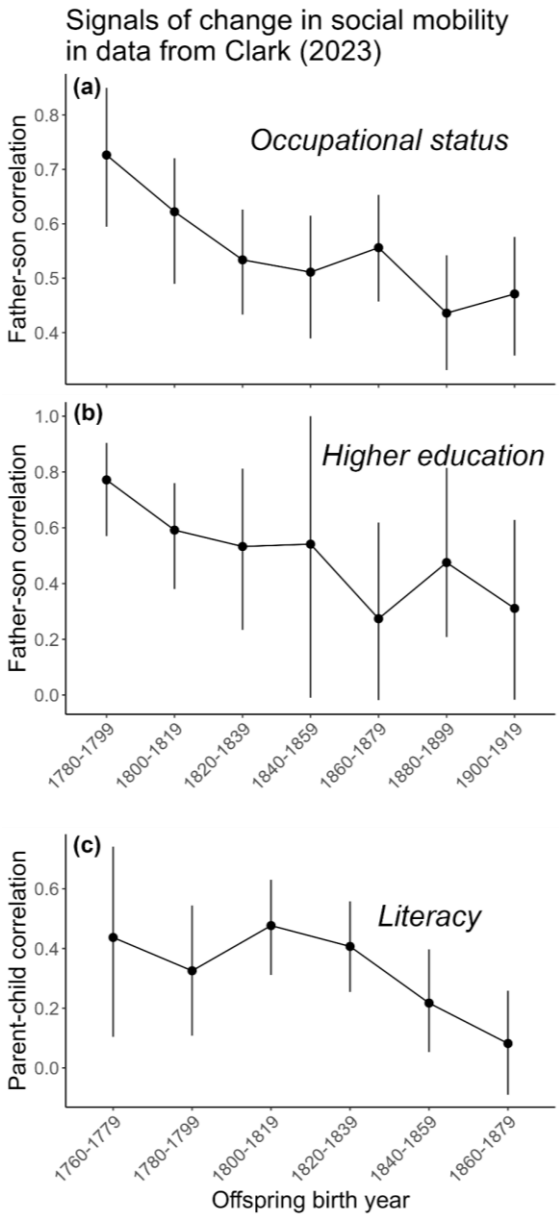

Figure 4. Signals of change in social mobility in data from Clark (2023). Parent-offspring correlations in multiple status measures generally decrease over the measured time period in (30)'s data, in contrast to claims of stagnant social mobility made in the original paper. **(a)** Occupational status 1780-1919; **(b)** higher education 1780-1919; **(c)** literacy 1760-1879. To mitigate pseudoreplication, we calculated correlations using one pair from each surname. Shown are average correlations (95% CI) across 500 bootstrap iterations of correlation estimation. **Fig. S13** shows two complementary analyses estimating correlations either without accounting for pseudoreplication, or using percentile ranks—both result in similar trends.

Confounding in genomic studies is downplayed. Human geneticists acknowledge residual confounding as an unsolved problem. But in practice, researchers face incentives to include genetic associations that are vulnerable to confounding in their analyses. Consider, for example, variation in polygenic (or "complex") traits, including virtually all behavioral traits, in which genetic contributions to trait variation are largely coming from numerous genetic variants with small individual effects. Researchers often wish to leverage weaker and weaker genetic associations to capture these highly polygenic signals. At the same time, confounding tends to be more severe the more weakly associated variants are considered (62, 68). Thus, in the pursuit of understanding polygenic effects, researchers may face a tradeoff between explaining a smaller part of the phenomenon under study in a causally rigorous way, versus explaining a seemingly larger part at the price of unknown biases introduced by confounding.

An example of this tradeoff lies in genetic trait prediction with so-called “polygenic scores” (69). Predictors based on more variants, including weakly-associated ones, often achieve higher prediction accuracy (62, 70). When predictive performance is prioritized, such predictors may be preferred by researchers, despite their higher susceptibility to confounding. Subsequent “consumers” like clinicians, researchers, policymakers, and the general public may then assume these polygenic scores capture strictly direct genetic effects, with the possibility of confounding rarely acknowledged.

Confounding is often assumed to be completely remedied by current methods, despite consistent evidence to the contrary (60–62, 64, 65, 67, 68, 71–73). Sometimes, methods to estimate genetic parameters grow in popularity even after they are shown to be susceptible to confounding, with this susceptibility rarely mentioned as a caveat (see, e.g., discussions in 60, 66, 67, 71, 74, 75).

In other cases, potential confounding is downplayed or obscured. As one example, consider the reporting of evidence for genetic effects from standard GWASs versus family studies. Family studies identify genotype-trait associations within, instead of among, families. This approach mitigates many sources of confounding (62, 64, 76). Family studies have yielded estimates of substantially weaker genetic effects on behavior or social outcomes (51, 62, 65, 67, 77–80). Reporting practices have in some cases obscured this point by asserting that there is a true genetic effect based on evidence from family studies, and emphasizing the magnitude of effects as estimated in a standard GWAS (37). Such reporting choices mislead by presenting signals susceptible to confounding as measures of genetic causality.

Demanding the highest standards of rigor for claims about genetics of social outcomes. Clark (30) uses a mechanistic model that is invalid for the data. It does not account for pervasive, uneven pseudoreplication. It suggests a new metric (of social mobility), neither justifying the validity of the new metric nor discussing the large discrepancy between the resulting findings and those based on other well-established metrics. These three core flaws can arise in any scientific field. But the study of the genetics underlying social outcomes, with its fraught history and heightened potential for misinterpretation and misappropriation, demands the highest standard of scientific rigor and scholarship. We are concerned that a publishing culture that rewards sensationalism may instead promote a decline in standards.

Conclusion. The study of heredity in humans has long been plagued by failures to address the implications of confounding between genetic and non-genetic sources of variation. Even today, when the inherent limitations of observational data are well appreciated, some studies continue to ignore, downplay, or even leverage such confounding in advancing claims about an outsized role for genetics. The failure to reckon with confounding encumbers scientific progress and can fuel the misappropriation of genetics research.

Acknowledgements

We thank Kate Antonovics, Mark Borrello, Raj Chetty, Graham Coop, Doc Edge, Sasha Gusev, Kelley Harris, Mark Kirkpatrick, Magnus Nordborg, Molly Przeworski, Sohini Ramachandran, Noah Rosenberg, James Schmitz, Elizabeth Thompson, and anonymous reviewers for comments on the manuscript and helpful discussions. We thank Gregory Clark for providing us with a corrected version of the occupational status data. The work was funded by NSF DBI-2010892 to J.W.B and NIH grant R35GM151108 to A.H.

Data availability. All code for reproducing the analyses herein is available at https://github.com/harpak-lab/Clark2023.

References

1. S. F. Galton, *Hereditary Genius: An Inquiry Into Its Laws and Consequences* (D. Appleton, 1869) https:/doi.org/10.1037/13474-000.

2. B. Mehler, Hereditarianism. *The Wiley Blackwell Encyclopedia of Race, Ethnicity, and Nationalism*, 1–3 (2015).

3. S. Wright, Statistical Methods in Biology. *J. Am. Stat. Assoc.* **26**, 155–163 (1931).

4. L. L. Cavalli-Sforza, M. W. Feldman, Models for cultural inheritance I. Group mean and within group variation. *Theor. Popul. Biol.* **4**, 42–55 (1973).

5. L. L. Cavalli-Sforza, M. W. Feldman, Cultural versus biological inheritance: phenotypic transmission from parents to children. (A theory of the effect of parental phenotypes on children's phenotypes). *Am. J. Hum. Genet.* **25**, 618–637 (1973).

6. D. C. Rao, N. E. Morton, S. Yee, Analysis of family resemblance. II. A linear model for familial correlation. *Am. J. Hum. Genet.* **26**, 331–359 (1974).

7. D. C. Rao, N. E. Morton, S. Yee, Resolution of cultural and biological inheritance by path analysis. *Am. J. Hum. Genet.* **28**, 228–242 (1976).

8. L. L. Cavalli-Sforza, M. W. Feldman, The Evolution of Continuous Variation. III. Joint Transmission of Genotype, Phenotype and Environment. *Genetics* **90**, 391–425 (1978).

9. J. Rice, C. R. Cloninger, T. Reich, Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of the unitary models. *Am. J. Hum. Genet.* **30**, 618–643 (1978).

10. C. R. Cloninger, J. Rice, T. Reich, Multifactorial inheritance with cultural transmission and assortative mating. II. a general model of combined polygenic and cultural inheritance. *Am. J. Hum. Genet.* **31**, 176–198 (1979).

11. C. R. Cloninger, J. Rice, T. Reich, Multifactorial inheritance with cultural transmission and assortative mating. III. Family structure and the analysis of separation experiments. *Am. J. Hum. Genet.* **31**, 366–388 (1979).

12. J. Rice, C. R. Cloninger, T. Reich, Analysis of behavioral traits in the presence of cultural transmission and assortative mating: Applications to IQ and SES. *Behav. Genet.* **10**, 73–92 (1980).

13. R. C. Lewontin, S. P. R. Rose, L. J. Kamin, *Not in Our Genes: Biology, Ideology, and Human Nature* (Pantheon Books, 1984).

14. B. J. Vilhjálmsson, M. Nordborg, The nature of confounding in genome-wide association studies. *Nat. Rev. Genet.* **14**, 1–2 (2013).

15. M. W. Feldman, F. B. Christiansen, S. P. Otto, Gene-culture co-evolution: teaching, learning, and correlations between relatives. *Isr. J. Ecol. Evol.* **59**, 72–91 (2013).

16. G. Solon, Theoretical models of inequality transmission across multiple generations. *Res. Soc. Stratif. Mobil.* **35**, 13–18 (2014).

17. N. Barton, J. Hermisson, M. Nordborg, Why structure matters. *Elife* **8** (2019).

18. R. Uchiyama, R. Spicer, M. Muthukrishna, Cultural evolution of genetic heritability. *Behav. Brain Sci.* **45**, e152 (2022).

19. M. D. Collado, I. Ortuño-Ortín, J. Stuhler, Estimating Intergenerational and Assortative Processes in Extended Family Data. *Rev. Econ. Stud.* **90**, 1195–1227 (2023).

20. A. F. Herzig, C. Noûs, A. S. Pierre, H. Perdry, A model for co-occurrent assortative mating and vertical cultural transmission and its impact on measures of genetic associations. *bioRxiv*, 2023.04.08.536101 (2023).

21. R. C. Lewontin, Annotation: the analysis of variance and the analysis of causes. *Am. J. Hum. Genet.* **26**, 400–411 (1974).

22. M. W. Feldman, R. C. Lewontin, The heritability hang-up. *Science* **190**, 1163–1168 (1975).

23. R. C. Bailey, Hereditarian scientific fallacies. *Genetica* **99**, 125–133 (1997).

24. N. A. Holtzman, Genetics and social class. *J. Epidemiol. Community Health* **56**, 529–535 (2002).

25. M. W. Feldman, S. Ramachandran, Missing compared to what? Revisiting heritability, genes and culture. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373** (2018).

26. A. I. Young, S. Benonisdottir, M. Przeworski, A. Kong, Deconstructing the sources of genotype-phenotype associations in humans. *Science* **365**, 1396–1400 (2019).

27. H. Shen, M. W. Feldman, Drowning in shallow causality. *Behav. Brain Sci.* **46**, e199 (2023).

28. A. S. Goldberger, Heritability. *Economica* **46**, 327–347 (1979).

29. A. Panofsky, Misbehaving science: Controversy and the development of behavior genetics. **321** (2014).

30. G. Clark, The inheritance of social status: England, 1600 to 2022. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2300926120 (2023).

31. R. A. Fisher, The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans. R. Soc. Edinb.* **52**, 399–433 (1918).

32. A. Gimelfarb, A general linear model for the genotypic covariance between relatives under assortative mating. *J. Math. Biol.* **13**, 209–226 (1981).

33. G. Clark, *The Son Also Rises* (Princeton University Press, 2014) https:/doi.org/10.1515/9781400851096 (October 13, 2023).

34. G. Clark, As a hereditarian, I strongly support economic redistribution. *Quillette* (2023) (August 21, 2023).

35. K. P. Harden, *The Genetic Lottery: Why DNA Matters for Social Equality* (Princeton University Press, 2021).

36. A. Panofsky, Biology meets public policyThe Genetic Lottery: Why DNA Matters for Social Equality Kathryn Paige Harden Princeton University Press, 2021. 312 pp. *Science* **373**, 1449 (2021).

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

38. G. Coop, M. Przeworski, Luck, lottery, or legacy? The problem of confounding. A reply to Harden. *Evolution* **76**, 2464–2468 (2022).

39. J. Fletcher, Backdoor to a dead end: A review essay. *Popul. Dev. Rev.* **48**, 253–258 (2022).

40. F. J. Odling-Smee, “Niche-constructing phenotypes” in *The Role of Behavior in Evolution , (pp*, H. C. Plotkin, Ed. (The MIT Press, viii, 1988), pp. 73–132.

41. Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).

42. E. Uffelmann, *et al.*, Genome-wide association studies. *Nature Reviews Methods Primers* **1**, 1–21 (2021).

43. P.-R. Loh, *et al.*, Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).

44. M. Hällsten, F. T. Pfeffer, Grand Advantage: Family Wealth and Grandchildren’s Educational Achievement in Sweden. *Am. Sociol. Rev.* **82**, 328–360 (2017).

45. ,*Rising Inequality, Schools, and Children’s Life Chances* (Russell Sage Foundation, 2011).

46. E. Karagiannaki, The effect of parental wealth on children’s outcomes in early adulthood. *J. Econ. Inequality* **15**, 217–243 (2017).

47. R. K. Q. Akee, W. E. Copeland, G. Keeler, A. Angold, E. J. Costello, Parents’ Incomes and Children's Outcomes: A Quasi-experiment Using Transfer Payments from Casino Profits. *Am. Econ. J. Appl. Econ.* **2**, 86–115 (2010).

48. K. Cooper, K. Stewart, Does Household Income Affect children’s Outcomes? A Systematic Review of the Evidence. *Child Indic. Res.* **14**, 981–1005 (2021).

49. M. Feldman, L. L. Cavalli-Sforza, *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton University Press, 1981) (October 3, 2023).

50. E. Turkheimer, Three Laws of Behavior Genetics and What They Mean. *Curr. Dir. Psychol. Sci.* **9**, 160–164 (2000).

51. A. I. Young, *et al.*, Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat. Genet.* **50**, 1304–1310 (2018).

52. L. Yengo, *et al.*, Imprint of assortative mating on the human genome. *Nat Hum Behav* **2**, 948–954 (2018).

53. S. H. Hurlbert, Pseudoreplication and the Design of Ecological Field Experiments. *Ecol. Monogr.* **54**, 187–211 (1984).

54. A. Miles, *Social Mobility in Nineteenth- and Early Twentieth-Century England* (Palgrave Macmillan UK, 1999) https:/doi.org/10.1057/9780230373211 (March 7, 2024).

55. J. J. Lee, The heritability and persistence of social class in England. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2309250120 (2023).

56. C. Cosh, Colby Cosh: Is socioeconomic status hereditary? *National Post* (2023) (September 14, 2023).

57. G. N. Marks, Has Cognitive Ability Become More Important for Education and the Labor Market? A Comparison of the Project Talent and 1979 National Longitudinal Survey of Youth Cohorts. *J Intell* **11** (2023).

58. M. Feldman, Echoes of the past: hereditarianism and A Troublesome Inheritance. *PLoS Genet.* **10**, e1004817 (2014).

59. A. S. Young, Estimation of indirect genetic effects and heritability under assortative mating. *bioRxiv*, 2023.07.10.548458 (2023).

60. J. J. Berg, *et al.*, Reduced signal for polygenic adaptation of height in UK Biobank. *Elife* **8** (2019).

61. M. Sohail, *et al.*, Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife* **8**, e39702 (2019).

62. H. Mostafavi, *et al.*, Variable prediction accuracy of polygenic scores within an ancestry group. *Elife* **9** (2020).

63. E. S. Lander, N. J. Schork, Genetic dissection of complex traits. *Science* **265**, 2037–2048 (1994).

64. C. Veller, G. Coop, Interpreting population and family-based genome-wide association studies in the presence of confounding. *bioRxiv*, 2023.02.26.530052 (2023).

65. A. Okbay, *et al.*, Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat. Genet.* **54**, 437–449 (2022).

66. M. N. Meyer, *et al.*, Wrestling with Social and Behavioral Genomics: Risks, Potential Benefits, and Ethical Responsibility. *Hastings Cent. Rep.* **53 Suppl 1**, S2–S49 (2023).

67. M. G. Nivard, *et al.*, More than nature and nurture, indirect genetic effects on children's academic achievement are consequences of dynastic social processes. *Nat Hum Behav* (2024) https:/doi.org/10.1038/s41562-023-01796-2.

68. A. J. Aw, J. McRae, E. Rahmani, Y. S. Song, Highly parameterized polygenic scores tend to overfit to population stratification via random effects. *bioRxiv*, 2024.01.27.577589 (2024).

69. A. Torkamani, N. E. Wineinger, E. J. Topol, The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 581–590 (2018).

70. B. J. Vilhjálmsson, *et al.*, Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).

71. R. Border, *et al.*, Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* **378**, 754–761 (2022).

72. G. Sella, N. H. Barton, Thinking About the Evolution of Complex Traits in the Era of Genome-Wide Association Studies. *Annu. Rev. Genomics Hum. Genet.* **20**, 461–493 (2019).

73. A. I. Young, *et al.*, Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nat. Genet.* **54**, 897–905 (2022).

74. S. Zabad, A. P. Ragsdale, R. Sun, Y. Li, S. Gravel, Assumptions about frequency-dependent architectures of complex traits bias measures of functional enrichment. *Genet. Epidemiol.* **45**, 621–632 (2021).

75. N. LaPierre, B. Fu, S. Turnbull, E. Eskin, S. Sankararaman, Leveraging family data to design Mendelian randomization that is provably robust to population stratification. *Genome Res.* **33**, 1032–1041 (2023).

76. R. S. Spielman, R. E. McGinnis, W. J. Ewens, Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**, 506–516 (1993).

77. J. J. Lee, *et al.*, Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).

78. S. Trejo, B. W. Domingue, Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses. *Biodemography Soc. Biol.* **64**, 187–215 (2018).

79. S. Selzam, *et al.*, Comparing Within- and Between-Family Polygenic Score Prediction. *Am. J. Hum. Genet.* **105**, 351–363 (2019).

80. L. J. Howe, *et al.*, Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).