



To Know and Know Not: Human Genetics and the Limits to Inference

John Benning¹

[LINK TO ABSTRACT](#)

One of the research programs in my lab is experimental evolution of *E. coli* to better understand the factors facilitating versus constraining adaptation to new environments. Recently, we started seeing some strange data—populations grown from the same clonal isolate, in the same treatment, on the same 96-well plate, began growing very differently, and the differences tracked position on the plate. We are still troubleshooting the source of this spatial gradient in growth, but whatever the underlying force—a malfunctioning heat source, or a gradient in oxygen availability, perhaps—the effect is strong enough to reverse which *E. coli* genotype ‘wins’ in a novel, high-stress environment.

Across even six inches of plastic, environmental differences confound inference of the cause of phenotypic differentiation and interact with genetic differences in unpredictable ways. Such is the ordinary condition of research in evolutionary biology and genetics. Environmental factors are structured in ways nobody anticipates. They correlate and interact with what you are trying to measure, and they lead researchers to make inferences about the causes of phenotypic variation that are coherent yet wrong. We guard against false inference with experiments randomizing genotypes across the environments they experience, which allow a clearer picture of the causal pathways among genes, environments, and phenotypes. Looking back through centuries of observational data on English families unfortunately offers a much muddier view.

¹ Department of Ecology & Evolutionary Biology, Cornell University.

The Identifiability Problem

This complex tangle between genes and environment is the pattern I see in my own work across plants, insects, and microbes. It sets my priors for evolutionary genetics: large environmental effects on phenotypes, unpredictable gene-by-environment interactions, and strong confounding of environmental and genetic similarity in non-experimental populations. So when I read about observational, highly structured data that purportedly “decisively reject[s]” social transmission, that suggest “it is direct genetic factors that shape social outcomes”, and that show that money spent on education is “largely a social waste” (Clark 2026, 61, 9, 218), I am skeptical. Such claims about the overwhelming primacy of genetics in shaping phenotypic variation demand strong evidence. Greg Clark’s *To Have and Have Not: The Determinants of Social Status in England, 1600–2026*, abbreviated here as THHN, has not provided that evidence.

As coauthors and I have argued (Benning et al. 2025), the central problem is the one that most people would intuit: though parents transmit genes, they also transmit wealth, property, social connections, educational expectations, language, and a surname and social network that open or close doors. Because we can often expect relatives to share such non-genetic inheritance in rough proportion to how closely related they are, genes and non-genetic factors are transmitted in tandem along the same pedigree. A model with one transmitted latent factor and an assortment parameter will fit correlations among relatives whether that factor is genetic, socio-cultural, or a mixture. The fit tells us nothing about the relative contribution of genetic and non-genetic inheritance to an individual’s phenotype. As decades of work have shown us (see references in Benning et al. 2025), inference as to the source of resemblance between relatives falls apart when genetic similarity cannot be separated from environmental similarity.

Thus, the parameter the book calls b^2 is in fact t^2 , total transmissibility, and m is the spousal correlation in that bundle of transmitted factors, genetic and non-genetic (Benning et al. 2025, section 2a and Supp. Note 1). Our interpretation does not refute the headline THHN result that persistence of status metrics is very high. But with the current analyses, and perhaps with any analyses of data where the confounding between genetic and non-genetic transmission is so complete, we cannot properly adjudicate what underlies the persistence of social status correlations among relatives. This difficulty is illustrated clearly in THHN, where much is made of the equality of parent-child and sibling correlations in social status metrics, and the reader is told that such a pattern “decisively rejects” a model of social transmission (Clark 2026, 61). This equality does indeed reject the only social model the book specifies, where the transmitted quantity is a family-level attribute that siblings share exactly, and thus siblings are always expected to resemble each

other more than they resemble their parents. But under alternative social models with decades of precedent, where siblings are allowed some variation in the non-genetic endowments they pass on (Cloninger et al. 1979), equality of parent-child and sibling correlations would be unremarkable.

The simplicity of the Fisher model that forms the backbone of THHN is appealing. But evolutionary genetics has long since realized such models are inaccurate for analyses of human populations. There have been many efforts over the past fifty years to better analyze observational data like those of the Families of England (FOE). Cavalli-Sforza and Feldman (1973) provided the model for cultural transmission taking a form similar to genetic transmission and said that such inheritance “may be difficult to distinguish” from the biological kind and that, in the parent-offspring case, it is “almost completely confounded” with it. Their proposed remedy was adoption data. The tradition that followed (through Rao et al. 1976 to Cloninger et al. 1979) exists precisely to estimate both genetic and non-genetic channels jointly, and to establish which data allow for their unique identification. It is for a clear reason that these models are more complex than the Fisher model used in THHN: after decades of work it is apparent that the dynamics of nature and nurture are not captured by any ‘genes only’ or ‘environment only’ model. You will find almost no modern biologist proposing such models—and trust me, we would be thrilled if understanding phenotypic variation were so simple.

Nowhere in THHN is a model containing both a genetic and a non-genetic transmission path fitted to the FOE data. The Fisher model is employed throughout, and a battery of ad hoc tests of environmental influence is offered in place of estimation.

As Benning et al. (2025) showed, core analyses in THHN are not robust to contestable assumptions (Supplementary Note 4), severe pseudoreplication (Supplementary Note 6), and reasonable alternative analyses (Supplementary Notes 4 and 7).² Each “environmental test” in Chapters 7–14 picks at a single strand of the tangled bundle of non-genetic attributes transmitted through families. Many of these attributes, like class position, social network, and culture, do not ‘split’ among siblings in a way that makes them useful instruments (Chapter 8), and they persist despite the death of a parent (Chapter 12) or a change of geography (Chapter 10). When class, network connections, wealth, property, and education are correlated and provide redundant insurance against household shocks, their individual contributions will often look insignificant under narrow, single-factor tests of a purely cultural model.

Even the cleaner instruments that the dataset has available offer limited pur-

² Clark published an EHES working paper (Clark 2025) with responses to some of the critiques in an earlier version of Benning et al. (2025). Those responses do not assuage my concerns regarding the analyses in THHN, and do not address the core problem of identifiability.

chase on quantifying the effects of non-genetic transmission on variation in social status metrics. Chapter 13 shows that, for families with above-average wealth, wealth shocks as delivered by variation in sibship size (i.e., individual inheritance decreases with sibship size) have no discernible effect on those siblings' descendants by the fourth generation; Clark (2026, 208) concludes that this shows that “the main mechanism of wealth transmission ... is not the actual physical transfer of wealth”. First, showing that a shock to inherited wealth is not detectable four generations later does not demonstrate that wealth does not influence status. Second, it is unsurprising that evidence for such a shock dissipates. Sibship size is a noisy proxy for individual inheritance, and, based on my calculations from the book's figures,³ the resulting shock plausibly accounts for perhaps a few percentage points of the variance in wealth at death in the analyzed, wealthy cohort. This test does not speak to the arguably more relevant question: does variance in wealth among families explain variance in status among families? It tells us nothing about how the difference in inherited wealth between the son of a miner and the son of a magistrate influences the trajectories of those men's lives.

Moving Forward

Newer data and better analyses have begun to resolve the sources of phenotypic differentiation in humans, often using molecular genetic data. Consistently, we see the estimated influence of direct genetic effects on many socio-cultural phenotypes fall as proper family and environmental controls are included. Alexander I. Young, Michael L. Frigge, Daniel F. Gudbjartsson, Gudmar Thorleifsson, et al. (2018) introduced relatedness disequilibrium regression, exploiting the random variation in relatedness generated by Mendelian segregation, and estimated the heritability of educational attainment at 17 percent. Recently, genomic analyses of more than 20,000 Norwegian parent-offspring trios estimated b^2 of educational attainment at 31 percent before controlling for family and environmental confounds, and $b^2 = 17$ percent with such controls in place (Corfield et al. 2026). Using a dataset that, like FOE, lacks molecular genetic data, María Dolores Collado, Ignacio Ortuño-Ortín, and Jan Stuhler (2023) separated a genetic from a cultural latent factor across more than 100 kinship types in Swedish registers. They attributed 7 percent of the variance in years of schooling to genes and ~34 percent to culture. They tested two variants of the simple genetic model proposed by Clark and found

³ Reading from the figures and text (Clark 2026, 206), an 8 percent decline in wealth for each sibling as sibship moves from four to five implies a coefficient on $\ln(\text{sibship})$ of -0.37 at the second generation. Reading the family-size distribution from Figure 13.2 gives a standard deviation of $\ln(\text{sibship})$ of 0.68, so the shock contributes roughly 0.25 log points of spread in $\ln(\text{wealth})$. I could not find the distribution of $\ln(\text{wealth})$ for this cohort reported anywhere in the book, so the share of variance cannot be estimated exactly; across a plausible range of spread it is a low single-digit percentage.

that they performed markedly worse than their baseline model (Collado et al. 2023, 1219).⁴

Contrast the estimates from the studies above with the estimate in THHN Table 5.4, where b^2 of years of education is estimated at 0.64. For the one outcome where the comparison is most direct, the book's estimate is four to nine times that reported from analyses built to properly identify genetic vs. non-genetic inheritance. I do not see why the estimates of genetic effects in THHN should carry more weight than studies specifically attempting to address the well-known identifiability problem. We should not lower the bar for evidence of the primacy of genetics simply because the data that could potentially separate genetic and non-genetic channels of inheritance in certain data sets are not available or have not been analyzed. The field spent decades building analytical designs to separate these channels, and a model that cannot separate them is no longer where the interesting questions are.

I do not claim, nor did my coauthors and I claim in Benning et al. (2025), to know to what extent social status is mediated by non-genetic influences. I am claiming that THHN's analyses of correlations among relatives in an extremely structured population cannot separate genetic and non-genetic transmission any better than Galton could. I am claiming that a patchwork of ad hoc tests—each pulling singly at one strand and each testing a model of purely cultural transmission that no one would suggest—is weak evidence that non-genetic transmission matters little for social status. THHN demonstrates that the Families of England database is an incredible data set, assortative mating is strong, and social status is indeed very persistent throughout family trees. But these analyses do not allow us to determine the genetic provenance of that persistence.

Could these data better separate channels of transmission if analyzed differently, taking advantage of information like the correlations in status of individuals with their in-laws? Perhaps, and the FOE data set certainly is a fascinating one to play with. I propose a series of open letters here in *Econ Journal Watch*. We can analyze the same open data set using robust, reproducible methods that allow everyone to follow along, and attempt to pin down as precisely as we can the sources of variation in social status among these families of England. Perhaps we'd learn something new.

⁴ In THHN, Clark claims this discrepancy is due to Collado et al. (2023)'s failure to adjust for measurement error. He refits his three-parameter genetic model with measurement error incorporated and judges it to fit "as well as the more elaborate Collado et al. model," a comparison he makes "just by visual inspection" (Clark 2026, 78). No fit statistics are compared between models.

References

- Benning, John W., Jedidiah Carlson, Olivia S. Smith, Ruth G. Shaw, and Arbel Harpak.** 2025. Confounding Fuels Misinterpretation in Human Genetics. *Proceedings of the Royal Society B: Biological Sciences* 292(2038): 20251615. [Link](#)
- Cavalli-Sforza, Luigi Luca and Marcus W. Feldman.** 1973. Cultural Versus Biological Inheritance: Phenotypic Transmission from Parents to Children (A Theory of the Effect of Parental Phenotypes on Children's Phenotypes). *The American Journal of Human Genetics* 25(6): 618–637. [Link](#)
- Clark, Gregory.** 2025. Hereditarian Fallacies? Inheritance of Social Status in England, 1600–2022. EHES Working Paper 285. European Historical Economics Society. [Link](#)
- Clark, Gregory.** 2026. To Have and Have Not: The Determinants of Social Status in England, 1600–2026. Unpublished manuscript. [Link](#)
- Cloninger, C. Robert, John Rice, and Theodore Reich.** 1979. Multifactorial Inheritance with Cultural Transmission and Assortative Mating. II. A General Model of Combined Polygenic and Cultural Inheritance. *The American Journal of Human Genetics* 31(2): 176–198. [Link](#)
- Collado, María Dolores, Ignacio Ortuño-Ortín, and Jan Stuhler.** 2023. Estimating Intergenerational and Assortative Processes in Extended Family Data. *The Review of Economic Studies* 90: 1195–1227. [Link](#)
- Corfield, Elizabeth C., Alexey A. Shadrin, Oleksandr Frei, Zillur Rahman, et al.** 2026. Family Genetic Designs in MoBa Provide Insights into Health and Functioning. *Nature*: 1–9. [Link](#)
- Rao, Dabeeru C., Newton E. Morton, and Shirley Yee.** 1976. Resolution of Cultural and Biological Inheritance by Path Analysis. *The American Journal of Human Genetics* 28(3): 228–242. [Link](#)
- Young, Alexander I., Michael L. Frigge, Daniel F. Gudbjartsson, Gudmar Thorleifsson, et al.** 2018. Relatedness Disequilibrium Regression Estimates Heritability Without Environmental Bias. *Nature Genetics* 50: 1304–1310. [Link](#)

About the Author



John Benning is an assistant professor in the Department of Ecology and Evolutionary Biology at Cornell University, where his lab studies how populations adapt, or fail to adapt, to novel and changing environments. The work combines field experiments, experimental evolution, and population genomics across plants, insects, and microbes. With coauthors he recently examined how confounding fuels misinterpretation in human genetics (*Proceedings of the Royal Society B*, 2025), taking Clark's 2023 PNAS paper as one of two case studies. See benninglab.org for more about his work. His email address is jbenning@cornell.edu.

[Go to Gregory Clark Symposium](#)
[Go to archive of Watchpad section](#)
[Go to September 2026 issue](#)