Int. J. Epidemiol. Advance Access published April 27, 2006

Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology
doi:10.1093/ije/dyl065

Commentary: The analysis of variance and the social complexities of genetic causation

Jeremy Freese[1,2]

Scientific articles published 30 years earlier can be interesting to revisit for various reasons. Lewontin's classic article on the analysis of variance in human behavioural genetics warrants continued attention for perhaps the worst of them: the article makes several correct observations that continue to remain under-appreciated in some research and much discussion about the causal role of genes in human outcomes.[1] The lucidity of Lewontin's arguments has historically proven no match for the allure of overly simple characterizations of outcomes as being x% due to genes and $(1 - x)$% not due to genes.[2] Moreover, Lewontin's main points speak beyond questions about genetics and could even be said to prefigure the best parts of more recent complaints about regression analysis as a tool for causal inference in observational studies.[3]

The problem is manifested in the social statistics analogue of an optical illusion: when one says x% of variance in some outcome (say, depression) is attributable to genes, one appears to be making a statement about the ultimate *nature* of that outcome. Heritability estimates are estimates of population parameters, however, meaning that they depend crucially on how genotypes and environments are distributed in the studied population. When calculations are made on oddly composed samples and with little information about either genotypes or environments, appropriate interpretation of precise heritability estimates can become downright mysterious. Likewise, discussions of the 'effect' of a risk factor on a disease are typically based on estimates of population-specific 'average causal effects'.[4] Whenever genetic effects vary across environments and the effects of risk factors vary across persons, simple answers follow only when questions are posed in terms of aggregate characterizations of particular populations, and such answers may change radically as populations or their aggregate circumstances change.

Lewontin recognized that population-dependent estimates are a poor substitute for actually understanding the specific action and interactions of causes. Lewontin erred, however, in declaring heritability estimates to be 'useless'. Estimates of population parameters are perfectly meaningful and possibly useful when properly recognized for what they are.[5] As importantly, the cumulation of high heritability outcomes from many studies have together had the genuinely useful consequence of making headway against various unfortunate, if often well-intentioned, resistances to contemplating the genetic contribution to behaviour that remain all too pervasive in some quarters of social science. This has provided a better foundation for social science to engage data and research with genotypic measures as they become increasingly available.[6]

Even so, commonsense understandings of causes and effects are quickly overwhelmed by the potential complexities of genetic causation in real human affairs, in ways that go well beyond the points Lewontin emphasizes regarding gene–environment interactions. Lewontin articulates the lessons of his article by reference to graphs of 'norms of reaction' by which the phenotypic outcomes associated with different genotypes vary over some environmental characteristic. (Prototypic is to imagine different plant genotypes producing different average heights depending on the sandiness of the surrounding soil.) Norms of reaction work far better for hypothetical plants than actual people, especially once one begins to consider the ways genes can systematically affect selection into environments. Environments are themselves outcomes that often ultimately implicate the causal import of genetic variation, like a plant whose genotype affects the sandiness of the soil in which it is planted. The ubiquitous tendency to oppose 'genes' and 'environments' becomes increasingly tired the more we understand the pervasive ways in which environments themselves are outcomes that depend on the interaction of psychological characteristics that are indisputably genetically influenced with larger social processes. Usual estimates of the heritability of disease, for example, reflect not only the influence of genes on 'biological' disease processes but also whatever influence genes have on behaviours and social selection processes that are associated with exposure or susceptibility.

Additionally, when thinking about health, understanding the causal influence of genes on environments must also be considered in the face of advances in treatment and social inequalities in resources to promote and protect their health. For example, cognitive and personality traits known to be partially inherited may affect propensities to adhere to complex self-administered treatment regimens.[7] Such traits become additionally pertinent to outcomes only to whatever extent such treatments are discovered and are effective. Moreover, these same traits may affect the ease with which one can obtain access to such treatments or to external sources of assistance in identifying and handling adherence problems, which may be both also influenced by government policies.[8] In addition, partially inherited psychological traits linked to educational attainment seem probably implicated in who takes up new innovations in health and who responds to new health information, which suggests that the heritabilities of health attributable to these processes will be associated with rates of

[1] Robert Wood Johnson Scholars in Health Policy Research Program, Harvard University, Cambridge, MA, USA.

[2] Department of Sociology, University of Wisconsin-Madison, Madison, WI, USA.

E-mail: jfreese@ssc.wisc.edu

progress in generating new knowledge about disease and the public diffusion of that knowledge.[9,10]

In other words, medical advances alter the environmental terrain of disease activity, and introduce and alter gene–environment interactions for disease outcomes in whatever ways effective treatment ends up being selective. Potentially, as advances decrease heritabilities in health outcomes attributable to proximate 'biological' causes of disease, it can increase heritabilities through the roles genes play in the determination of psychological traits and sociological circumstances that affect the extent to which individuals gain the full benefit of these advances. (For that matter, treatment advances are not themselves necessarily genotypically neutral, as their pursuit is strongly influenced by expectations about the number and remunerative capacity of beneficiaries.)[11] Consequently, the observed heritability of disease outcomes in human populations depends not only on the joint distribution of genotypes and phenotypes but also on distributions of knowledge and resources.

For that matter, as humans gain information not only about the genetics of disease but also about the contents of their own genotypes, complexities of gene–environment causality emerge that non-human genetics has never needed to contemplate. Among other things, the causal effects of genes no longer require phenotypic expression of those genes, for information about genetic risk of some condition can initiate possibly successful preventive action. Once again, the propensity to learn about such risks and to respond with preventive action may depend on heritable cognitive and personality traits and the influence of such traits on social circumstances. This would imply heritability in behaviours that effectively suppress other kinds of heritable variation.

A central message of Lewontin's article was that the specific articulation of causal relations is not much served by accounting-style exercises of variance decomposition. The complicated ways genes can be causally relevant for health outcomes further underscores this point. At the same time, it also emphasizes the necessary complement of psychological and sociological inquiry for a complete understanding of the causality of genes. Those interested in gene–environment interactions understand well the need this implies for better measures of environment and better understanding of the specific causal pathways by which environmental pathogens operate.[12] Beyond this, however, a full reckoning of genetic causation will also require much improvement in our understanding of the unfolding and historically contingent relationships between partially inherited phenotypic traits and the social processes underlying effective disease prevention and treatment.

References

[1] Lewontin RC. The analysis of variance and the analysis of causes. *Am J Hum Genet* 1974;**26:**400–411. (Reprinted *Int J Epidemiol* doi:10.1093/ije/dyl062.)

[2] Alford J, Funk C, Hibbing JR. Are political orientations genetically transmitted? *Am Polit Sci Rev* 2005;**99:**153–67.

[3] Berk RA. Regression analysis: a constructive critique. Thousand Oaks, CA: Sage; 2004.

[4] Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health* 2000;**21:**121–45.

[5] Freese J, Li JCA, Wade LD. The potential relevances of biology to social inquiry. *Annu Rev Sociol* 2003;**29:**233–56.

[6] Caspi A, McClay J, Moffitt T *et al*. Evidence that the cycle of violence in maltreated children depends on genotype. *Science* 2002;**297:**851–54.

[7] Goldman D, Smith J. Can patient self-management help explain the SES health gradient? *Proc Natl Acad Sci USA* 2002;**99:**10929–34.

[8] Lutfey K, Freese J. Toward some fundamentals of fundamental causality: socioeconomic status and health in the routine clinic visit for diabetes. *Am J Sociol* 2005;**110:**1326–72.

[9] Glied S, Lleras-Muney A. *Health inequality, education, and medical innovation. NBER Working Paper No. 9738*. Cambridge, MA: National Bureau of Economic Research, 2003.

[10] Meara E. *Why is health related to socioeconomic status? The case of pregnancy and low birth weight. NBER Working Paper No. 8231*. Cambridge, MA: National Bureau of Economic Research, 2001.

[11] Kremer M, Glennerster R. *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*. Princeton, NJ: Princeton University Press, 2004.

[12] Moffitt TE, Caspi A, Rutter M. Strategy for investigating interactions between measured genes and measured environments. *Arch Gen Psychiatry* 2005;**62:**473–81.