Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2006;**35**:520–525
doi:10.1093/ije/dyl062

REPRINTS AND REFLECTIONS

The analysis of variance and the analysis of causes

R C LEWONTIN*

This issue of the *American Journal of Human Genetics* contains two articles by Newton Morton and his colleagues[1,2] that provide a detailed analytic critique of various estimates of heritability and components of variance for human phenotypes. They make especially illuminating remarks on the problems of partitioning variances and covariances among groups such as social classes and races. The most important point of all, at least from the standpoint of the practical, social, and political applications of human population genetics, occurs at the conclusion of the first paper[1] in which Morton points out explicitly the chief programmatic fallacy committed by those who argue so strongly for the importance of heritability measures for human traits. The fallacy is that a knowledge of the heritability of some trait in a population provides an index of the efficacy of environmental or clinical intervention in altering the trait either in individuals or in the population as a whole. This fallacy, sometimes propagated even by geneticists, who should know better, arises from the confusion between the technical meaning of heritablility and the everyday meaning of the word. A trait can have a heritability of 1.0 in a population at some time, yet this could be completely altered in the future by a simple environmental change. If this were not the case, 'inborn errors of metabolism' would be forever incurable, which is patently untrue. But the misunderstanding about the relationship between heritability and phenotypic plasticity is not simply the result of an ignorance of genetics on the part of psychologists and electronic engineers. It arises from the entire system of analysis of causes through linear models, embodied in the analysis of variance and covariance and in path analysis. It is indeed ironic that while Morton and his colleagues dispute the erroneous programmatic conclusions that are drawn from the analysis of human phenotypic variation, they nevertheless rely heavily for their analytic techniques on the very linear models that are responsible for the confusion.

I would like to look rather closely at the problem of the analysis of causes in human genetics and to try to understand how the underlying model of this analysis moulds our view of the real world. I will begin by saying some very obvious and elementary things about causes, but I will come thereby to some very annoying conclusions.

Discrimination of causes and analysis of causes

We must first separate two quite distinct problems about causation that are discussed by Morton. One is to discriminate which of two alternative and mutually exclusive causes lies at the basis of some observed phenotype. In particular, it is the purpose of *segregation analysis* to attempt to distinguish those individuals who owe their phenotypic deviation to their homozygosity for rare deleterious gene alleles from those whose phenotypic peculiarity arises from the interaction of environment with genotypes that are drawn from the normal array of segregating genes of minor effect. This is the old problem of distinguishing major gene effects from 'polygenic' effects. I do not want to take up here the question of whether such a clear distinction can be made or whether the spectrum of gene effects and gene frequencies is such that we cannot find a clear dividing line between the two cases. The evidence at present is ambiguous, but at least *in principle* it may be possible to discriminate two etiological groups, and whether such groups exist for any particular human disorder is a matter for empirical research. It is possible, although not necessary, that the form of clinical or environmental intervention required to correct a disorder arising from homozygosity for a single rare recessive allele (the classical 'inborn error of metabolism') may be different from that required for the 'polygenic' class. Moreover, for the purposes of genetic counselling, the risk of future offspring being affected, will be different if a family is segregating for a rare recessive than if it is not. Thus, the discrimination between two *alternative* causes of a human disorder is worth making if it can be done.

The second problem of causation is quite different. It is the problem of the *analysis* into separate elements of a number of causes that are interacting to produce a single result. In particular, it is the problem of analyzing into separate components the interaction between environment and genotype in the determination of phenotype. Here, far from trying to discriminate individuals into two distinct and mutually exclusive etiological groups, we recognize that all individuals owe their phenotype to the biochemical activity of their genes in a unique sequence of environments and to developmental events that may occur subsequent to, although dependent upon, the initial action of the genes. The analysis of interacting causes is fundamentally a different concept from the discrimination of alternative causes. The difficulties in the early history of genetics embodied in the pseudo-question of 'nature versus nurture' arose precisely because of the confusion between these two problems in causation. It was supposed that the

Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA.

* First published Am J Hum Genet 1974 26:400-411. Reprinted with permission.

phenotype of an individual could be the result of *either* environment *or* genotype, whereas we understand the phenotype to be the result of *both*. This confusion has persisted into modern genetics with the concept of the phenocopy, which is supposed to be an environmentally caused phenotypic deviation, as opposed to a mutant which is genetically caused. But, of course, both 'mutant' and 'phenocopy' result from a unique interaction of gene and environment. If they are etiologically separable, it is not by a line that separates environmental from genetic causation but by a line that separates two kinds of genetic basis: a single gene with major effect or many genes each with small effect. That is the message of the work by Waddington[3] and Rendel[4] on canalization.

Quantitative analysis of causes

If an event results from the joint operation of a number of causative chains and if these causes 'interact' in any generally accepted meaning of the word, it becomes conceptually impossible to assign quantitative values to the causes of that *individual event*. Only if the causes are utterly independent could we do so. For example, if two men lay bricks to build a wall, we may quite fairly measure their contribution by counting the number laid by each; but if one mixes the mortar and the other lays the bricks, it would be absurd to measure their relative quantitative contributions by measuring the volumes of bricks and of mortar. It is obviously even more absurd to say what proportion of a plant's height is owed to the fertilizer it received and what proportion to the water, or to ascribe so many inches of a man's height to his genes and so many to his environment. But this obvious absurdity appears to frustrate the universally acknowledged program of Cartesian science to analyze the complex world of appearances into an articulation of causal mechanisms. In the case of genetics, it appears to prevent our asking about the relative importance of genes and environment in the determination of phenotype. The solution offered to this dilemma, a solution that has been accepted in a great variety of natural and social scientific practice, has been the *analysis of variation*. That is, if we cannot ask how much of an individual's height is the result of its genes and how much a result of its environment, we will ask what proportion of the deviation of height from the population mean can be ascribed to deviation of environment from the average environment and how much to the deviation of this genetic value from the mean genetic value. This is the famous linear model of the analysis of variance which can be written as

$$Y - \mu_Y = (G - \mu_Y) + (E - \mu_Y) + (GE) + e, \qquad (1)$$

where μ_Y is the mean score of all individuals in the population; Y is the score of the individual in question; G is the average score of all individuals with the same genotype as the one in question; E is the average score of all individuals with the same environment as the one in question; GE, the genotype-environment interaction, is that part of the average deviation of individuals sharing the same environment and genotype that cannot be ascribed to the simple sum of the separate environmental and genotypic deviations; and e takes into account any individual deviation not already consciously accounted for, and assumed to be random over all individuals (measurement error, developmental noise, etc.).

I have written this well known linear model in a slightly different way than it is usually displayed in order to emphasize two of its properties that are well known to statisticians. First, the environmental and genotypic effects are in units of *phenotype*. We are not actually assessing how much variation in environment or genotype exists, but only how much perturbation of phenotype has been the outcome of average difference in environment. The analysis in equation (1) is completely *tautological*, since it is framed entirely in terms of phenotype and both sides of the equation must balance by the definitions of GE and e. To turn equation (1) into a contingent one relating actual values of environmental variables such as temperature to phenotypic score, we would need functions of the form

$$(E - \mu_Y) = f(T - \mu_T) \qquad (2)$$

and

$$GE = h\left[\left(g - \mu_g\right), (T - \mu_T)\right], \qquad (3)$$

where g and T are measured on a genetic and a temperature scale rather than on a scale of phenotype. Thus, the linear model, equation (1), makes it impossible to know whether the environmental deviation $(E - \mu_Y)$ is small because there are no variations in actual environment or because the particular genotype is insensitive to the environmental deviations, which themselves may be quite considerable. From the standpoint of the tautological analysis of equation (1), this distinction is irrelevant, but as we shall see, it is supremely relevant for those questions that are of real importance in our science.

Second, equation (1) contains population means at two levels. One level is the grand mean phenotype μ_Y and the other is the set of so-called 'marginal' genotypic and environmental means, E and G. These, it must be remembered, are the *mean* for a given environment averaged over all genotypes in the population and the *mean* for a given genotype averaged over all environments.

But since the analysis is a function of these phenotypic means, it will, in general, give a different result if the means are different. That is, the linear model is a *local analysis*. It gives a result that depends upon the actual distribution of genotypes and environments in the particular population sampled. Therefore, the result of the analysis has a historical (i.e., spatiotemporal) limitation and is not in general a statement about *functional* relations. So, the genetic variance for a character in a population may be very small because the functional relationship between gene action and the character is weak for any conceivable genotype or it may be small simply because the population is homozygous for those loci that are of strong functional significance for the trait. The analysis of variation cannot distinguish between these alternatives even though for most purposes in human genetics we wish to do so.

What has happened in attempting to solve the problem of the analysis of causes by using the analysis of variation is that a totally different object has been substituted as the object of investigation, almost without noticing it. The new object of study, the deviation of phenotypic value from the mean, is not the same as the phenotypic value itself; and the tautological analysis of that deviation is not the same as the analysis of causes. In fact, the analysis of variation throws out the baby

with the bath water. It is both too specific in that it is spatiotemporally restricted in its outcome and too general in that it confounds different causative schemes in the same outcome. Only in a very special case, to which I shall refer below, can the analysis of variation be placed in a one-to-one correspondence to the analysis of causes.

Norm of reaction

The real object of study, both for programmatic and theoretical purposes, is the relation between genotype, environment, and phenotype. This is expressed in the *norm of reaction*, which is a table of correspondence between phenotype, on the one hand, and genotype-environment combinations on the other. The relations between phenotype and genotype and between phenotype and environment are many-many-relations, no single phenotype corresponding to a unique genotype and vice versa.

In order to clarify the relation between the two objects of study (i.e. the norm of reaction and the analysis of variance, which analyses something quite different), let us consider the simplified norms of reaction shown in figures 1*a–h*. We assume that there is a single well-ordered environmental variable *E*, say temperature, and a scale of phenotypic measurement *P*. Each line is the norm of reaction, the relationship of phenotype to environment, for a particular hypothetical genotype (G_1 *or* G_2).

The first thing to observe is that in every case the phenotype is sensitive to differences in both environment and genotype. That is, each genotype reacts to changing environment, and in no case are the two genotypes identical in their reactions. Thus in any usual sense of the word, both genotypes and environment are *causes* of phenotypic differences and are necessary objects of our study.

Figure 1*a* is in one sense the most general, for if environment extends uniformly over the entire range and if the two genotypes are equally frequent, there is an overall effect of genotype (G_1 being on the average superior to G_2) and an overall effect of environment (phenotype gets smaller on the average with increasing temperature). Nevertheless, the genotypes cross so that neither is always superior.

Figure 1*b* shows an overall effect of environment, since both genotypes have a positive slope; but there is no overall effect of genotype, since the two genotypes would have exactly the same *mean* phenotype if all environments wereconsidered equally. There is no *a priori* way from Figure 1*b* of ranking the two genotypes. However, if because of particular circumstances the distribution of environments were heavily weighted toward the lower temperatures, then G_1 would be consistently superior to G_2; an analysis of variance would show a strong effect of genotype as well as of environment, but very little genotype-environment interaction. Thus the analysis of variance would reflect the particular environmental circumstances and give a completely incorrect picture of the general relationship between cause and effect here, where there is overall no effect of genotype but a strong genotype-environment interaction.

Figure 1*c* is the complementary case to that shown in Figure 1*b*. In figure 1*c* there is no overall effect of environment, but G_1 is clearly superior to G_2 overall. In this case a strong environmental component of variance will appear, however, if either one of the genotypes should predominate in the population. So the historical events that mould the genotypic distribution of a population will have an effect on the judgment, from the analysis of variance, of the importance of environment.

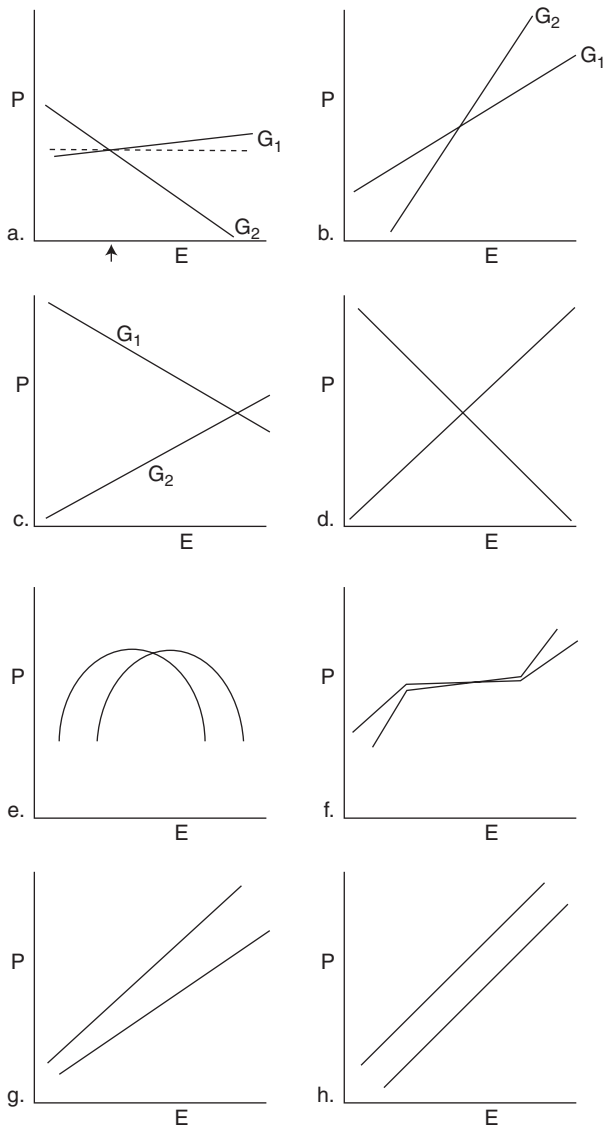

Figure 1 *a–h*, Examples of different forms of reaction norms. In each case the phenotype (*P*) is plotted as a function of environment (*E*) for different genotypes (G_1, G_2).

The overall lack of genetic effect in Figure 1*b* and of environmental effect in Figure 1*c* can both appear in a trait like that shown in Figure 1*a*, which overall has both effects if the distribution of environments or of genotypes is asymmetric. Thus if environments are distributed around the middle in Figure 1*a*, there will *appear* to be no average effect of genotype, while if the population is appropriately weighted toward an excess of G_1, the average phenotype across environments will be constant as shown by the dashed line. Here real overall effects are obscured because of spatiotemporal events, and the analysis of variance fails to reveal significant overall differences.

These last considerations lead to two extremely important points about the analysis of variance. First, although equation (1) appears to isolate distinct causes of variation into separate elements, it does not do so because the amount of *environmental* variance that appears depends upon the *genotypic* distribution, while the amount of *genetic* variance depends upon the *environmental* distribution. Thus the appearance of the separation of causes is a pure illusion. Second, because the linear model appears as a sum of variation from different causes, it is sometimes erroneously supposed that removing one of the sources of variation will reduce the total variance. So, the meaning of the genetic variance is sometimes given as ''the amount of variation that would be left if the environment were held constant,' and the environmental variance is described as 'the amount of variance that would remain if all the genetic variation were removed,' an erroneous explanation offered by Jensen[5], for example. Suppose that the norms of reaction were as in figure 1*a* and a unimodal distribution of environments were centered near the middle, with a roughly equal mixture of the two genotypes. Now suppose we fix the environment. What will happen to the total variance? That depends on which environment we fix upon. If we choose an environment about 1 SD or more to the right of the mean, there would actually be an *increase* in the total variance because the difference between genotypes is much greater in that environment than on the average over the original distribution. Conversely, suppose we fix the genotype. If we chose G_2 to be our pure strain, then, again, we would *increase* the total variance because we had chosen the more environmentally plastic genotype. The apparent absurdity that removing a source of variance actually increases the total variance is a consequence of the fact that the linear model does not really effect a separation of causes of variation and that it is a purely local description with no predictive reliability. Without knowing the norms of reaction, the present distribution of environments, and the present distribution of genotypes, and without then specifying which environments and which genotypes are to be eliminated or fixed, it is impossible to predict whether the total variation would be increased, decreased, or remain unchanged by environmental or genetic changes.

In Figure 1*d* there is no overall effect of either genotype or environment, but both can obviously appear in a particular population in a particular environmental range as discussed above.

The case shown in 1*e* has been chosen to illustrate a common situation for enzyme activity, a parabolic relation between phenotype and environment. Here genotypes are displaced horizontally (have different temperature optima). Neither geontype is superior overall, nor is there any general monotone environmental trend for either genotype. But for any distribution of environments except a perfectly symmetrical one, there will appear a component of variance for genotypic effect. Moreover, if the temperature distribution is largely to either side of the crossover point between these two genotypes, there will be very large components of variance for both genotype and environment and a vanishingly small interaction component; yet over the total range of environments exactly the opposite is true!

Figure 1*e* also shows a second important phenomenon, that of differential phenotypic sensitivity in different environmental ranges. At intermediate temperatures there is less difference between genotypes and less difference between the effect of environments than at more extreme temperatures. This phenomenon of canalisation, is more generally visualized in Figure 1*f*. Over a range of intermediate phenotypes there is little effect of either genotype or environment, while outside this zone of canalisation, phenotype is sensitive to both[4]. The zone of canalisation corresponds to the range of environments that have been historically the most common in the species, but in new environments much greater variance appears. Figure 1*f* bears directly on the characteristic of the analysis of variance that all effects are measured in phenotypic units. The transformations, equations (2) and (3), that express the relationship between the phenotypic deviations ascribable to genotype or environment and the actual values of the genotypes or environmental variables are not simple linear proportionalities. The sensitivity of phenotype to both environment and genotype is a function of the particular range of environments and genotypes. For the programmatic purposes of human genetics, one needs to know more than the components of variation in the historical range of environments.

Figures 1*a-f* are meant to illustrate how the analysis of variance will give a completely erroneous picture of the causative relations among genotype, environment, and phenotype because of the particular distribution of genotypes and environments in a given population at a given time picks out relations from the array of reaction norms that are necessarily atypical of the entire spectrum of causative relations. Of course it may be objected that any sample from nature can never give exactly the same result as examining the universe. But such an objection misses the point. In normal sampling procedures, we take care to get a representative or unbiased sample of the universe of interest and to use unbiased sample estimates of the parameters we care about. But there is no question of sampling here, and the relation of sample to universe in statistical procedures is not the same as the relation of variation in spatiotemporally defined populations to causal and functional variation summed up in the norm of reaction. The relative size of genotypic and environmental components of variance estimated in any natural population reflect in a complex way four underlying relationships: (1) the actual functional relations embodied in the norm of reaction; (2) the actual distribution of genotype frequencies—a product of long-time historical forces like natural selection, mutation, migration, and breeding structure—which changes over periods much longer than a generation; (3) the actual structure of the environments in which the population finds itself, a structure that may change very rapidly indeed, especially for human populations; and (4) any differences among genotypes that may cause a biased distribution of genotypes among environments. These differences may be behavioural (for instance, a heat-sensitive genotype may seek cooler habitats), or it may result from other individuals using the genotype as an indicator for differential treatment, since that treatment is part of environment. A causal pathway may go from tryptophane metabolism to melanin deposition to skin color to hiring discrimination to lower income, but equation (1) would simple indicate heritability for 'economic success'. The effects of historical forces and immediate environment are inextricably bound up in the

outcome of variance analysis which thus is not a tool for the elucidation of functional biological relations.

Effect of additivity

There is one circumstance in which the analysis of variance can, in fact, estimate functional relationships. This is illustrated exactly in figure 1*h* and approximately in figure 1*g*. In these cases there is perfect or nearly perfect additivity between genotypic and environmental effects so that the differences among genotypes are the same in all environments and the differences between environments are the same for all genotypes. Then the historical and immediate circumstances that alter genotypic and environmental distributions are irrelevant. It is not surprising that the assumption of additivity is so often made, since this assumption is necessary to make the analysis of variance anything more than a local description.

The assumption of additivity is imported into analyses by four routes. First, it is thought that in the absence of any evidence, additivity is *a priori* the simplest hypothesis and additive models are dictated by Occam's razor. The argument comes from a general Cartesian world view that things can be broken down into parts without losing any essential information and that in any complex interaction of causes, main effects will almost always explain most of what we see while interactions will tend to be of a smaller order of importance. But this is a pure *a priori* prejudice. Dynamic systems in an early state in their evolution will show rather large main effects of the forces acting to drive them, but as they approach equilibrium the main effects disappear and interactions predominate. For example, that is what happens to additive genetic variance under selection. Exactly how such considerations apply to genotype and environment is not clear.

Second, it is suggested that additivity is a first approximation to a complex situation, and the results obtained with an additive scheme are then a first approximation to the truth. This argument is made by analogy with the expansion of mathematical functions by Taylor's series. But this argument is self-defeating since the justification for expanding a complex system in a power series and considering only the first-order terms is precisely that one is interested in the behavior of the system in the neighborhood of the point of expansion. Such an analysis is a local analysis only, and the analysis of variance is an analysis in the neighborhood of the population mean only. By justifying additivity on this ground, the whole issue of the global application of the result is sidestepped.

Third, it is argued that if an analysis of variance is carried out and the genotype-environment interaction turns out to be small, the assumption of additivity is justified. As in the second argument, there is some circularity. As the discussion of the previous section showed, the usual outcome of an analysis of variance in a particular population in a restricted range of environments is to underestimate severely the amount of interaction between the factors that occur over the whole range.

Finally, additivity or near additivity may be assumed without offering any justification because it suits a predetermined end. Such is the source of figure 1*g*. It is the hypothetical norm of reaction for IQ taken from Jensen[5] . It purports to show the relation between environmental 'richness' and IQ for different genotypes. While there is not a scintilla of evidence to support such a picture, it has the convenient properties that superior and inferior genotypes in one environment maintain that relation in all environments, and that as environment is 'enriched,' the genetic variance (and therefore the heritability) increases. This is meant to take care of those foolish egalitarians who think that spending money and energy on schools generally will iron out the inequalities in society.

Evidence on actual norms of reaction is very hard to come by. In man, measurements of reaction norms for complex traits are impossible because the same genotype cannot be tested in a variety of environments. Even in experimental animals and plants where genotypes can be replicated by inbreeding experiments or cloning, very little work has been done to characterize these norms for the genotypes that occur in natural populations and for traits of consequence to the species. The classic work of Clausen *et al*[6]. on ecotypes of plants shows very considerable non-additivity of the types illustrated in Figures 1*a–d*.

As an example of what has been done in animals, Figure 2 has been drawn from the data of Dobzhansky and Spassky[7] on larval viability in *Drosophila pseudoobscura*. Each line is the reaction norm for larval viability at three different temperatures for a fourth-chromosome homozygote, where the chromosomes have been sampled from a natural population. As the figure shows, a few genotypes are of uniformly poor viability, probably corresponding to homozygosity for a single deleterious gene of strong effect. However, most genotypes are variable in their expression, and there is a great deal of genotype-environment interaction with curves crossing each other and having quite different environmental sensitivities.

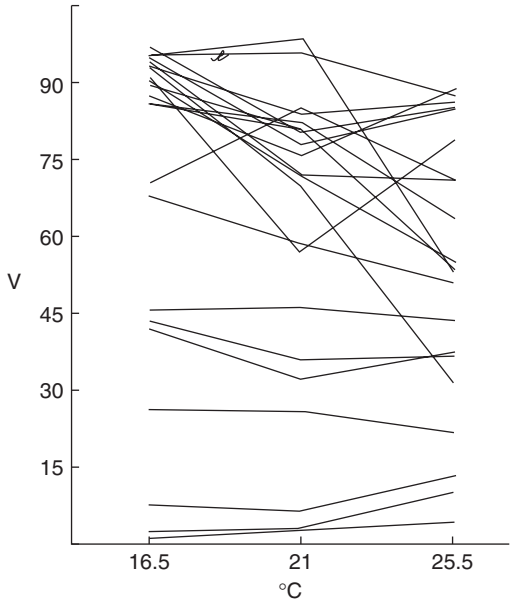

Figure 2 Actual reaction norms for viability of fourth chromosome homoyzgotes of *Drosophila pseudoobscura*. Data from Dobzhansky and Spassky [7].

Purpose of analysis

Just as the objects of analysis are different when we analyze causes and when we analyze variance, so the purposes of these analyses are different. The analysis of causes in human genetics is meant to provide us with the basic knowledge we require for correct schemes of environmental modification and intervention. Together with a knowledge of the relative frequencies of different human genotypes, a knowledge of norms of reaction can also predict the demographic and public health consequences of certain massive environmental changes. Analysis of variance can do neither of these because its results are a unique function of the present distribution of environment and genotypes.

The legitimate purposes of the analysis of variance in human genetics are to predict the rate at which selection may alter the genotypic composition of human populations and to reconstruct, in some cases, the past selective history of the species. Neither of these seems to be a pressing problem since both are academic. Changes in the genotypic composition of the species take place so slowly as compared to the extraordinary rate of human social and cultural evolution, that human activity and welfare are unlikely to depend upon such genetic change. The reconstruction of man's genetic past, while fascinating, is an activity of leisure rather than of necessity. At any rate, both these objectives require not simply the analysis into genetic and environmental components of variation, but require absolutely a finer analysis of genetic variance into its additive and nonadditive components. The simple analysis of variance is useless for these purposes and indeed it has no use at all. In view of the terrible mischief that has been done by confusing the spatiotemporally local analysis of variance with the global analysis of causes, I suggest that we stop the endless search for better methods of estimating useless quantities. There are plenty of real problems.

References

[1] Morton NE. Analysis of family resemblance. I. Introduction. Am J Hum Genet 1974;**26**:318–330.

[2] Rao DC, Morton NE, Yee S. Analysis of family resemblance. II. A linear model for familial correlation. *Am J Hum Genet* 1974;**26**:331–359.

[3] Waddington CH. Genetic assimilation of an acquired character. *Evolution* 1953;**7**:118–126.

[4] Rendel JM. Canalization of the acute phenotype of Drosophila. *Evolution* 1959;**13**:425–439.

[5] Jensen AR How much can we boost IQ and scholastic achievement? *Harvard Educ Rev* 1969;**39**:1–123.

[6] Clausen J, Keck DD, Heisey WM Experimental studies on the nature of species. I. Effects of varied environments on western North American plants. *Carnegie Inst Washington Publ* 1940;**520**:1–452.

[7] Dobzhansky T, Spassky B Genetics of natural populations. XI. Manifestation of genetic variants in *Drosophila pseudoobscura* in different environments. *Genetics* 1944;**29**:270–290.

Published by Oxford University Press on behalf of the International Epidemiological Association *International Journal of Epidemiology* 2006;**35:**525–527

doi:10.1093/ije/dyl064

Commentary: Heritability estimates—long past their sell-by date

Steven P R Rose

Heritability then

It might seem—it probably is—presumptuous for a neuroscientist to comment on a theoretical text in population genetics, especially when the paper in question is by one of the prominent figures in the field. However, it is relevant to recall the context in which Lewontin's 1974 article in *The American Journal of Human Genetics*[1] appeared. Symbolized by the publication, in 1969, of Arthur Jensen's article: *How much can we boost IQ and scholastic achievement?*,[2] there had been a resurgence of claims as to the heritability of human traits. Jensen had argued that, as IQ scores had a high heritability (~80%), it followed that the consistent difference in IQ scores between black and white citizens of the US was too great to be accounted for by 'environmental' factors. Instead, he concluded, the on average lower IQ of blacks compared with whites must reflect genetic differences between the two populations.

Jensen's contention raised a firestorm of political and scientific responses (e.g. Kamin,[3] Gould,[4] Rose *et al.*[5]). Some of these focussed on empirical inadequacy, others on the theoretical limitations of the IQ theory and of heritability calculations. It is with the latter two that Lewontin's article is concerned. His intent is, first to clarify common misconceptions over the meaning of the term, and second, to emphasize its inutility outside the very narrow range of circumstances for which it was originally derived. To summarize:

(i) Heritability is *not* a measure of the contributions of genes and environments to any *individual* phenotype, a fruitless enterprise as both are subsumed within the processes of development.

Department of Biological Sciences, The Open University, Milton Keynes MK7 6AA, UK.

E-mail: s.p.r.rose@open.ac.uk