The Reverend Dr L J Eaves
(Department of Genetics, University of Birmingham, Birmingham, B15 2TT)

Human Behavioural Genetics

Quantitative geneticists are usually concerned with the understanding of normal variation. Clinicians are necessarily concerned with abnormalities. There may, however, be some common ground in that understanding normal variation may help various practitioners to comprehend in a wider context the abnormalities with which they are frequently faced.

In trying to understand why people differ in their behaviour we are not so much concerned these days with whether or not such differences are partly inherited. Generations of carefully designed laboratory and field experiments in plants and animals have rarely found a trait for which there is no genetical variation at all (*see* e.g. Falconer 1960). This applies as much to behaviour as it does to other kinds of trait (*see* e.g. McClearn & De Fries 1973). All would show clear evidence that both genetical and environmental factors are important components of individual differences. It would be as naive to assume that man *must* be different as it would be to assume that there are no problems in showing that he is not.

Experimental organisms have three advantages: (1) We can control the mating system. (2) We can randomize genotypes over environments. (3) We can replicate genotypes in different environments. None of these is practicable with man. We have to accept the mating system and the family structure we are given and recognize that the power of any analysis we do is likely to be much reduced. There is, however, an intrinsic interest in studying the biological significance of our culture and mating system, and our family

structure. Perhaps studies of man come as close as possible to the study of a natural population in its environment. The very fact that we can recognize people and families makes this much easier.

The newcomer to human behaviour genetics has to overcome two mental blockages before he can get inside the subject. The first is very simple. He must remember that most individual phenotypic differences are continuous, that is, they resemble differences in height. They tend not to be the discontinuous differences such as those we associate with clinical genetics. To assume that genetics is about discontinuous phenotypic differences is to ignore much of genetics since about 1910. The manifold discontinuities we find at the DNA level can produce continuity at the phenotypic level for a particular inherited trait (*see* e.g. Mather & Jinks 1971).

The second barrier which has to be overcome relates to the tacit assumption that so-called 'social' factors (family background, social class, income &c.) may be used as indices of the quality of an individual's environment. They may just as well be the cultural expressions of his genes. Only very careful analyses of extensive family data can help us decide which is the most plausible explanation.

In the early part of this century, the particulate basis of mendelian inheritance and the continuity of much variation were thought to be incompatible. There were empirical laws, such as Galton's law of ancestral heredity, for explaining away the degree of similarity between relatives, but there was no theoretical substance nor adequate explanation of the empirical findings. The outstanding theoretical contribution of Fisher (1918) provided the missing substance and launched the branch of genetics known as quantitative, or biometrical, genetics. He showed that the variation for a so-called metrical trait (that is, continuously variable trait) could be due to the cumulative effect of many loci, each of small individual effect, and each obeying quite precisely the laws of particulate inheritance as they had been elucidated by Mendel. Previous research had already shown how genetic discontinuity, coupled with environmental factors, could produce continuous variation in the phenotype; how genes could have identical and cumulative effects, and how they segregated in F_2s from crosses between pure breeding strains. Subsequent research had shown how the genes producing continuous variation were borne on chromosomes along with the genes which contributed to the more marked discontinuities of phenotype associated with classical mendelian genetics (*see* e.g. Mather & Jinks 1971).

This early work, and the area of genetical research which stemmed from it, was mainly successful because it was able to base its predication not merely on empirical laws, such as those of the early biometricians, but on the regular properties of particulate inheritance which enable us to write precise mathematical models for the degrees of similarity and difference between relatives.

Any model which leads us to precise quantitative expectations for the similarities and differences between relatives can be tested by comparing observed data with the values we might predict from a theoretical model. Later we shall illustrate our approach for psychoticism, a trait which has some clinical significance. A baseline for most studies of variation and one of the simplest of all reasonable models for individual differences is one in which we assume both genetical and environmental factors contribute to variation, but there are certain restraints upon the kind of genetical system underlying the variation and on the kinds of environmental factors which might affect the trait. Thus, we might assume for the sake of prediction, that as far as the particular trait was concerned, individuals chose their mates at randon and that the gene action was entirely additive. The latter means in effect that we are assuming that at each and every locus which contributes to variation a heterozygote is intermediate in expression between its corresponding homozygotes.

As far as the environment is concerned we might assume, again until proved otherwise, that all the environmental variation is due to experiences quite unique to individuals, and not shared with any other members of their families.

We could now examine the precise implications of our assumptions for human similarities and differences and see if the pattern of similarities and differences we observe for human relationships agrees with our predictions. The simple set of assumptions we have made can even be tested crudely with data on monozygotic and dizygotic twins reared together, since we expect the following: (1) MZ individuals to be no more nor less variable than DZ twins; (2) the degree of similarity of MZ twins (as measured by the correlation coefficient) to be twice as great as that of DZ twins.

Only if both of these criteria are met can we legitimately proceed with the simple genetical interpretation of twin data. There are other ways of formulating our expectations which are used in practice and we would hope to work not only with twins because they enable us to test only the most restricted set of assumptions (*see* e.g. Jinks & Fulker 1970). Both these expectations can be subjected to a statistical test with a given set of data for twins. If a given set of data fails to match the simple expectations we have to reject our provisional model for variation and think again.

I have chosen only the simplest case and a very restricted experimental design, namely the twin study, because it is important to establish the methodological principle that the particulate theory of inheritance leads on to a variety of precise quantitative models for individual differences which can be tested with appropriate experimental designs. I have stated my view that a simple model, such as the simple additive-genetic random-mating specific environmental effects model we have outlined, is merely the baseline for studies of variation. It represents a 'noise-level' against which other, perhaps more interesting, factors have to be detected and more critical investigations conducted.

There is little of fundamental interest today in the repetition of studies which show a genetical basis of individual differences. Issues of greater biological and clinical interest remain to be examined. We may distinguish, in principle, three areas we would like to understand much more fully.

(1) *Genetical*

The mating system: Can we assume mating is random, or is it assortative? If like tends to marry like in a population, genetical differences between individuals are much more apparent, since alleles of like effect tend to become associated in the same individuals. The mating system may have a considerable effect on this rate of evolution of a population. We would not be surprised if man, like other animals, showed assortative mating for those aspects of the behavioural phenotype most closely related to fitness.

Non-additive genetical variation: Is there much of this and if so, what sort? The detailed analysis of many traits in organisms other than man has suggested that certain kinds of dominant and epistatic effects are associated with traits displaying a strong linear relationship to fitness. Thus alleles which increase fitness tend to be dominant with respect to their decreasing counterparts and genes related to fitness seem to show epistatic interaction of the duplicate type. This finding provides a tool for uncovering the relationships between a given continuously variable trait and fitness since a knowledge of genetic architecture can illuminate our understanding of the adaptive significance of a trait. In practice, it may be quite difficult to detect genetical non-additivity for organisms for which controlled breeding is impossible.

Gene frequencies and the number of loci: Many attempts have been made to explain variation by reference to a few polymorphic genes of reduced penetrance. For autosomal traits it is very difficult to estimate the number of loci and the allele frequencies with any precision, but under some circumstances we can ask whether there is any evidence that the increasing alleles differ in frequency, on average, from their decreasing counterparts; and given that this is not the case, to prove at least a notional figure for the number of loci. Such estimates are extremely unreliable and it is perhaps unwise to take them at all seriously since they depend on many assumptions of which we have only a very poor test with human data.

(2) *Environmental Factors*

Clearly it is not enough to know that environmental heterogeneity is responsible for behavioural variation. In the end we wish to know what environmental factors are important. Although a detailed analysis of the environmental factors may be difficult – and may even be premature – we can identify three broad areas of environmental variation which can help to focus our attention more critically:

(*a*) *Unreliability of measurement:* In practice unreliability turns out to be one of the largest components of environmental variation. When we interview someone, or give a psychological test, we are doing it on a particular day at a particular time. We rarely get exactly the same test score twice. Indeed, such unreliability may be an inherent property of the behaviour which is being studied. Providing we have chosen an adequate scale of measurement, variation due to unreliability is usually included with estimates of the contribution of environmental factors in most experimental designs.

(*b*) *Specific environmental experiences:* During the course of an individual's development he may receive his own unique environmental influences which he shares with no other members of his family. There may be accidents, specific traumata, chance encounters and so on. They serve to make individuals within a family differ over and above any innate differences between them.

(*c*) *Shared environmental experiences:* Humans grow in families. In this respect they are far from ideal experimental organisms since they cannot be randomized over the range of environments. Any developmentally significant effects common to children in a family will tend to make them alike. Such effects might be social, educational, economic, maternal or paternal.

In distinguishing our three types of environmental factor we are making a distinction which is analytically tractable and practically important. Variation due to measurement error, for example, is not much use to the therapist, yet it may be a substantial component of individual differences

when these are assessed by personal interview or psychological test. Certain kinds of error may of course represent traits in themselves which could have clinical and biological significance. We might thus wish to devise ways of measuring and analysing genetically a subject's proneness to error. If there is a lot of sampling variation it might lead us to think we can do more by simple environmental manipulation than is in fact the case, by leading us to overestimate the importance of existing environmental factors.

Variation due to specific environmental experiences has usually to be detected over and above the noise of sampling error. If these differences are substantial then we have a good case for examining people's personal case histories rather than their family backgrounds for the environmental causes of the differences between them. The effects of shared environmental experiences, on the other hand, are the most difficult to detect in practice. Even if we have individuals reared in foster homes such factors can only be detected against the background of all other environmental factors. When we have individuals reared together we can only detect shared environmental influences against the background of a basic genetical expectation for the similarity between relatives.

There is, however, no real substitute for such studies. Attempts to identify actual environmental factors by measuring the social environment might be misleading because it is quite possible for the social environment to have a genetical component. This leads to a third area of valuable enquiry which has only recently been handled analytically. This comprises those factors which are neither purely genetic nor environmental but which represent the interaction or covariation of both.

(3) *Genotype-environmental Factors*

We may distinguish (*a*) genotype-environmental interaction and (*b*) genotype-environmental covariation. The first thing to say about both is that they do not undermine any analysis of variation but may complicate it.

(*a*) *Genotype-environmental interaction:* There is a popular confusion surrounding genotype-environmental interaction. We are not speaking here of the kind of physiological interaction (or co-action) of genotypic and environmental factors which enter into the development of any organism. We are considering the way in which genotypes differ in their responses to the same environmental change. Genotype-environmental interaction describes the situation in which the organism's sensitivity to environmental experiences will depend on the genotype of the individual concerned. The methods of quantitative genetics have dealt most effectively with genotype-environmental interaction in other organisms, and have shown how sensitivity to environmental differences is indeed genetically determined and may therefore be under the influence of natural selection. It is more difficult to analyse genotype-environmental interaction in man because it is not possible to expose the same genotype to a range of controlled environmental differences. There are, however, a few circumstances in which such analysis is possible. We may detect genotype-environmental interaction for example if the same genetical factors affect the trait directly and affect the sensitivity of the individual to environmental influences. We may also detect genotype-environmental interaction for any environmental factor with respect to which genetical differences are known to be distributed randomly in the population. In practice this may be very restricting since many if not most environmental variables by which we tend to describe populations are almost certainly not independent of genetical differences. This leads to genotype-environmental covariation.

(*b*) *Genotype-environmental covariations:* Jencks (1973) has described genotype-environmental covariation as a 'double advantage' phenomenon in which (if genotype-environmental covariation is positive) advantaged genotypes also tend to receive advantageous environments. Providing we are prepared to be more precise about the mechanism of underlying environmental variation we can, at least in principle, analyse genotype-environmental covariation by a natural extension of quantitative genetics. We may distinguish three kinds of genotype-environmental covariation (there may be others) due to: (i) individuals selecting their own environments; (ii) sibs' genotypes acting on part of their co-sibs' environments; (iii) parental genotypes forming part of the developmentally significant environment of their offspring's environments. The first, of course, is analytically intractable because genotype-environmental covariation arising for this reason is inseparable from genetical variation. The second and third may be detectable analytically and have considerable biological importance since the detection of such genotype-environmental covariation could imply that the human population is endowed with the raw genetical material for kinship selection. That is, the survival or sacrifice of one individual may promote the survival of a relative. They also imply that factors which we might tend, naively, to write off as environmental may have a genetical basis.

It is particularly important to recognize the possible covariation of genetical and environmental factors in behaviour, especially when we attempt to explain individual differences in

behaviour by reference to family background data, since we may mislead ourselves seriously about the ultimate cause of variation.

We must stress that many of the possible components of variation are not fundamentally intractable given large and appropriate experimental designs. Also, we recognize as a matter of biological fact that virtually any trait we measure in other organisms shows some genetical variation.

We may illustrate some of the principles by reference to a trait which may be of interest to clinicians, namely psychoticism. There is a school of thought which regards certain forms of psychopathology as extremes of a continuum. For example, Gottesman & Shields (1968) brought to light the possibility of treating schizophrenia as a threshold character for the purposes of genetical analysis, much as one might analyse diabetes.

Eysenck's P Scale, incorporated in personality questionnaires such as the PEN, and EPQ, is presented as one attempt to measure the hypothesized continuum of which certain forms of psychopathology may be one extreme.

So far we have only analysed data on MZ and DZ twins reared together, but we have collected data on parents and offspring, siblings and adopted relatives, and may in the future be able to present a test of the model I present here. Our sample sizes are fairly large by some standards; in one study we have 320 pairs of MZ and 258 pairs of DZ twins. An earlier study, using a less adequate scale, had rather more twins (Eaves & Eysenck, in preparation).

The individual's raw P score is obtained by adding up all the 'P' type responses to a series of items on the questionnaire. As such a scale has many undesirable properties we choose, for psychological and statistical reasons, to work with the square-roots of the raw scores. This assumes, in effect, that an individual's predisposition to psychoticism gives him a fixed (small) probability of responding to each of the items of the P scale in a 'psychotic' fashion. The apparent simplicity of the conclusions from our analysis of such scores suggests we were justified in our choice of scale.

In our analysis of the scores derived from the PQ we found little evidence of genotype-environmental interaction on this scale. This means that genetical and environmental effects simply add together to produce the phenotype with respect to psychoticism as we have measured it. There is no evidence that different genotypes differ in their sensitivity to environmental influences.

As far as the PQ was concerned we found that our baseline model for variation was adequate to explain the variation observed in male and female MZ and DZ twins. That is, we could find no evidence that their genes showed any non-additive effects; there was no indication of dominance or epistasis. Neither could we detect any genetical consequences of assortative mating. This is comforting, because the marital correlation for P is too small to produce a marked genetical effect. As far as the genetical system is concerned, therefore, our current model for P is the simplest there is, the genes acting in a simple additive fashion. We can find no evidence that increasing alleles are more or less frequent than decreasing alleles, and no evidence of directional non-additive effects. The only estimate we have of the number of loci suggests that they are many (perhaps more than five), but this figure could be so much in error as to be of little value.

Our picture of the genetical system, therefore, gives us little reason to suppose there is any strong linear relationship between psychoticism and fitness. That is, we would be surprised to find extreme psychotics any more or less fit than extreme normals. If we look anywhere for increased reproductive fitness we should perhaps turn to the middle of the range. An interesting sidelight at this point is the fact that this is the conclusion to which advocates of a single gene theory of schizophrenia have tended in order to explain the maintenance of a polymorphism for a trait that is obviously deleterious (*see* e.g. Slater & Cowie 1971).

We should not be misled by the apparent detail of these conclusions into accepting them incautiously. Non-additive genetical effects, which for this purpose include any directional inequality of allele frequencies, are notoriously difficult to detect even when they are quite marked. The main point, however, remains. We have little reason to suppose, at this stage, that the genetical system underlying variation in psychoticism is anything other than the simplest additive polygenic system.

In fact, only about 46% of the variation in psychoticism appeared to be genetically determined, so we are compelled to ask seriously 'What about the environment?'

As far as I can see from our data there is no marked effect of the shared family environment. The twin data give us no reason to suppose that the similarity between members of the same family is due to anything other than ordinary genetical similiarities.

If this is correct, clinicians may have to be very careful when they look at the family background to explain the origin of psychopathology. They may not be assessing the relevant environments, only the relevant genes.

All the environmental influences, therefore, amounting to around 54% of the total variation

in P, are due either to specific environmental experiences of individuals, or to unreliability of measurements. What are their relative contributions to the environmental variation? Fortunately, our psychological model for P is such that we can assess the expected sampling error quite easily. We find that about 75% of the environmental variation in P is due to sampling error, the remaining 25% being due to 'real' environmental influences specific to individuals in the same family.

This amounts to saying that about 11% of the total variation in psychoticism as we have measured it is due to environmental factors over and above the 'noise' level of genetical variation and sampling error. The place to look for these differences seems to be within the family, rather than between families, since they are specific to individuals rather than shared by families.

One question remains. Can we say anything about the covariation of genetical and environmental effects? The answer is 'not much'. Since there appears to be little shared environmental effect we might tentatively conclude that any genotype-environmental covariation arising because of maternal and paternal effects must be relatively small. Since the basic model fits the data on twins reared together we can be a little more confident that there is no genetic-environmental effect on siblings for the PQ, but there is an interesting caveat relating to the P scale of the earlier PEN. Here we found, for males only, a failure of the sample model of a kind which could be due to a competition effect between twins in the same family based on their respective genotypes. The data fitted remarkably well with what we would expect if the phenotype of one twin were exerting an environmental effect on the phenotype of the other in the pair, such that individuals predisposed towards psychoticism tended to make their co-twin more normal and vice versa. This finding, related as it is to one sex in one study, should not be taken too seriously but it does illustrate how a precise quantitative theory of individual differences can illuminate our understanding of behavioural variation and provide a stimulus to further enquiry.

REFERENCES

Falconer D S (1960) Introduction to Quantitative Genetics. Oliver & Boyd, Edinburgh

Fisher R A (1918) *Transactions of the Royal Society of Edinburgh* 52, 399–433

Gottesman I I & Shields J (1968) In: Progress in Human Behavior Genetics. Ed. S G Vandenberg. Johns Hopkins, Baltimore; p 64

Jencks C (1973) Inequality: A Reassessment of the Effect of Family and Schooling in America. Allen Lane, London

Jinks J L & Fulker D W (1970) *Psychological Bulletin* 73, 311–349

McClearn G E & De Fries J C (1973) Introduction to Behavior Genetics. W H Freeman, San Francisco

Mather K & Jinks J L (1971) Biometrical Genetics. Chapman & Hall, London

Slater E & Cowie V (1971) The Genetics of Mental Disorders. Oxford University Press, London