VOL. 73, NO. 5 MAY 1970

Psychological Bulletin

COMPARISON OF THE BIOMETRICAL GENETICAL, MAVA, AND CLASSICAL APPROACHES TO THE ANALYSIS OF HUMAN BEHAVIOR[1]

J. L. JINKS AND D. W. FULKER[2]

University of Birmingham, England

The techniques which can be used in the analysis of human behavior by the methods of biometrical genetics are described and compared with those of the Multiple Abstract Variance Analysis (MAVA), and other approaches. These techniques are applied to a number of personality and cognitive measures using published data. Underlying assumptions of the analyses used are discussed, and tests of significance for departure from them are demonstrated. Although data were often inadequate, the techniques provided new information on the gene action controlling the measures and on their evolution. The authors conclude that the outcome of the reanalyses indicates the unique value of the biometrical approach.

There are currently three alternative approaches to the genetical analysis of human twin and familial data. There is what might be termed the classical approach through correlations between relatives, culminating in the estimation of various ratios describing the relative importance of genetic and environmental influences on trait variation. This approach leads to ratios such as the *H* of Holzinger (1929), the *E* of Neel and Schull (1954), and the *HR* of Nichols (1965), each of which measures an aspect of the relative importance of heredity and environment. There is the more systematic and comprehensive approach of the Multiple Abstract Variance Analysis (MAVA) developed by Cattell (1960, 1965) leading to both the estimation of nature: nurture ratios, and an assessment of the importance of the correlation between genetic and environmental influences within the family as well as within the culture. This approach is open-ended and based on the comparison of within- and between-family variances of full- and half-sib families, as well as monozygotic and dizygotic twins. Finally, there is the biometrical genetical approach initiated by Fisher (1918), and extended and applied by Mather (1949), which includes the first two approaches as special cases, and attempts to go beyond them to an assessment of the kinds of gene action and mating system operating in the population. While the biometrical genetic approach in psychogenetics has been used almost exclusively and with considerable success in investigations with animals, it has not often been used in the investigation of human populations. It is the two previously mentioned methods, which were specially developed for this purpose, that have been employed with humans.

In view of the increasing awareness of the power of biometrical genetics among those working in the psychogenetic area, it would seem opportune to present an account of the application of some of its methods to human data. It is the purpose of this paper, therefore, to illustrate the biometrical approach by reference to data collected and analyzed by a number of other workers, and to underline its relationship to the other two approaches.

[1] Preparation of this paper was partially supported by grants from the Agricultural Research Council of Great Britain, and the United States Public Health Service, MH-08712, from the National Institute of Health.

[2] We are grateful to Professor P. L. Broadhurst for his help and encouragement during the preparation of this paper.

Requests for reprints should be sent to J. L. Jinks, Department of Genetics, or D. W. Fulker, Department of Psychology, University of Birmingham, P. O. Box 363, Birmingham, 15, England.

Within- and Between-Family Comparisons

Let us consider data collected from n families each consisting of m individuals. An analysis of variance to compare the variation within and between families leads to the following expectations:

Source	df	Expected MS
Between families	$n-1$	$\sigma^2_W + m\sigma^2_B$
Within families	$n(m-1)$	σ^2_W

where σ_W^2 is that part of the total variance (σ_T^2) due to differences within families, and σ_B^2 is the part due to differences between families.

Two approaches are current in biometrical genetics: One is to estimate σ_W^2 and σ_B^2 from the analysis of variance, and the other is to estimate directly from the data the average variation within-families $\bar{V}_F$, and the variation between-family means $V_{\bar{F}}$, where $\bar{V}_F$ = mean square within families, and $V_{\bar{F}} = 1/m$ (mean square) between families. Thus these two approaches are related as $\bar{V}_F = \hat{\sigma}_W^2$, and $V_{\bar{F}} = (\hat{\sigma}_W^2 + m\hat{\sigma}_B^2)/m$. Clearly where m, the family size, is very large $\bar{V}_F = \hat{\sigma}_W^2$, and $V_{\bar{F}} = \hat{\sigma}_B^2$.

In the MAVA approach, the σ^2s for specifying the variation within and between families are related to the above as follows:

$$\sigma_w^2 \text{ MAVA} = \bar{V}_F = \sigma_W^2$$
$$\sigma_b^2 \text{ MAVA} = mV_{\bar{F}} = \sigma_W^2 + m\sigma_B^2$$

that is, the σ^2s of MAVA are the corresponding mean squares derived from the analysis of variance, but the between-family item is neither a σ^2 nor a simple function of a σ^2.

Having estimated the Vs or σ^2s by one of the available systems, it is sometimes necessary to express them as proportions of the total variation. Indeed some of these proportions are the correlations of the classical approach. The total variation in the biometrical genetical approach is clearly $\sigma_T^2 = \sigma_W^2 + \sigma_B^2$. In the MAVA approach, on the other hand, σ_T^2 MAVA $= 2\sigma_W^2 + m\sigma_B^2$ which is not, of course, the total variation that we wish to partition among various sources. However, this can be accommodated in the special case of families of Size 2 ($m = 2$) discussed by Cattell because of σ_T^2 MAVA $= 2\sigma_W^2 + 2\sigma_B^2 = 2\sigma_T^2$.

Loehlin (1965) also gives an account of how the σ^2MAVA are defined and their relationship to σ^2 as conventionally defined. For an account of the use of $\bar{V}_F$ and $V_{\bar{F}}$, see Mather (1949).

When discussing expectations in terms of biometrical genetical models, in the general case, it is more convenient to work in terms of the σ^2s of standard analyses of variance, which are independent of experimental design and family size, rather than $\bar{V}_F$ and $V_{\bar{F}}$, or σ^2 MAVA, which are not. On the other hand, when these expectations are equated to observed values from a specific set of data in order to estimate the parameters of a model, the situation is reversed. Both expected and observed values must now be given in terms of $\bar{V}_F$ and $V_{\bar{F}}$, which are independent of one another, rather than the σ^2s, which are not, in order that procedures based on least squares may be used to estimate the parameters in the model.

Genotype-Environment Models

Following the decision to discuss the biometrical genetical models in relation to the variance components σ_W^2 and σ_B^2, it is necessary to consider the interpretation of their values in terms of a model which adequately describes the genetic and environmental contributions to the total variation.

Thus while $\sigma_T^2 = \sigma_W^2 + \sigma_B^2$, it also equals $\sigma_G^2 + \sigma_E^2 + f(G,E)$ where σ_G^2 = the genetic variation, σ_E^2 = the environmental variation, and $f(G,E)$ = some function of genotypic and environmental contributions, and may represent two distinct sources of variation. If genotypic and environmental contributions are correlated with respect to size and sign, $f(G,E) = 2$ cov (G,E). However, if environmental deviations depend for their absolute size (irrespective of sign) on the particular genotypic deviations paired with them, there is genotype-environment interaction and $f(G,E) = \sigma_{GE}^2$. While the effect of 2 cov (G,E) may be to increase or decrease apparent $\sigma_G^2 + \sigma_E^2$ the effect of σ_{GE}^2 will always be to increase it.

While those who use the classical approach through correlations are aware of the problems created by f(G,E), this approach has, as yet, been unable either to specify, detect, estimate, or correct for the effects of this source of variation.

MAVA and biometrical genetics, on the other hand, specifically recognize the effects of f(G,E), and seek to specify them in the models and to estimate these effects in the analysis. However, the approaches are quite different in a number of important respects. In the MAVA approach all possible contributors to f(G,E) are allowed for, but whether or not they are included in the expectation for any particular σ^2 MAVA is decided largely on a subjective assessment of the likelihood that they will contribute to it. When all possible sources of f(G,E) are included, the number of parameters in the models is raised to an almost unmanageable number and places the analysis beyond the reach of most bodies of data. In the biometrical approach the presence or absence of a particular item in the expectation is decided by the form of the expected contribution of genetic and environmental components according to invariable rules. However, the biometrical genetical approach does not stop at this point, for it poses the question whether or not the correlation or interaction items in the models are essential or redundant by means of a number of statistical tests (scaling tests) that specifically detect their presence.

As may be seen later from the reanalyses of data, these scaling tests allow us to suggest with some confidence that very simple genetical models are quite adequate to account for most of the data. Moreover, where we cannot make this assertion, we are in a position to judge the kind of extension of the model needed better to describe the data (see Example 2, Table 4), and to avoid the pitfall of suggesting the need to fit a complex model to data resulting from inadequate sampling of the population under study (see Examples 3 and 4, Table 4). We cannot emphasize too strongly the risk we feel is involved in fitting complex models like those of MAVA without first carrying out scaling tests and tests of the adequacy of the sampling to ensure that such models are either necessary or appropriate. To fit a complex model to inadequate data may well lead to completely unfounded conclusions. Together with the insight that biometrical genetics provides concerning gene action and the mating system of the population, it is in providing these tests that the chief value of the approach lies, as compared with its alternatives.

Tests for Genotype-Environment Interaction

Numerous tests for this purpose have been described for use with controlled plant and animal breeding programs, but none, so far, have been proposed for use with human data. This omission, however, poses no insuperable problems.

Suppose we have n families of monozygotic twins, such that the twins in Family 1 have scores t_{11} and t_{12}, those in family n being t_{n1} and t_{n2}, respectively. When the twins have been reared together, $t_{11} - t_{12} \ldots t_{n1} - t_{n2}$, each provides an estimate of the magnitude of environmental influences within families. If all twin pairs are affected to the same extent by the environmental influences within the family, then $t_{11} - t_{12} = t_{21} - t_{22} \ldots = t_{n1} - t_{n2}$, within sampling error. However, if twins in some families react differently from those in other families when exposed to the same environmental influences, or twins in some families are exposed to different environmental influences than in other families, then $t_{11} - t_{12} \neq \ldots \neq t_{n1} - t_{n2}$.

The sum of the twin scores $t_{11} + t_{12}$, $t_{21} + t_{22} \ldots t_{n1} + t_{n2}$, on the other hand, will differ if the twins belonging to different families have different genotypes, different family environments, or both. If there is any interaction between genotype, and within family environment, then we should find a correlation between the twin sums $t_{11} + t_{12}$ etc., and the twin differences $t_{11} - t_{12}$ etc. over the n families.

If the twins have been reared apart the same considerations will allow us to test for an interaction between genotype and environmental differences between families. Thus we can determine whether or not the assumption of independence of genotypic and environmental influences is valid, and hence, whether or not parameters of interaction between these influences should be included in the model.

The expected magnitude of the correlation between the difference between twins $(t_{11} - t_{12})$ etc. and their sums $(t_{11} + t_{12})$ etc., can be expressed in standard biometrical terms. Let us take the simplest of all situations where the twins differ by a single gene, A — a, in a random mating population in which the gene frequencies are equal. There are then three types of twin pairs AA, Aa, and aa occurring with the frequencies $\frac{1}{4}:\frac{1}{2}:\frac{1}{4}$. The genetic contributions to

TABLE 1

SCALED EFFECT OF A SINGLE LOCUS ON A CONTINUOUSLY VARYING TRAIT

Genotype	Scaled effect of genotype	
	With dominance	Without dominance
AA	mean + d	mean + d
Aa	mean + h	mean
aa	mean − d	mean − d

the mean scores of these three types of twins will be $+d$, h, and $-d$, respectively. The gene effects d, h, and $-d$ may be seen in Table 1. The item d clearly measures half the difference between the two homozygous genotypes AA and aa, while h measures the deviation of the heterozygote from the average of AA and aa. If there is no dominance then $h = 0$. A detailed account of this basic scaling procedure may be found elsewhere (Mather, 1949). Let us now consider a simple environmental difference e_1 acting within each family, and a single environmental difference e_2 acting between families. We will also allow the different genotypes to interact quite differently with the environmental differences. For example, the interaction between twins with genotype AA, $(+d)$ will interact with the within-family environmental difference (e_1) to the extent gd_1, and with the between-family environmental difference (e_2) to an extent gd_2. The values for heterozygous twins Aa will be gh_1 and gh_2, respectively. We then have 12 kinds of twin phenotypes measured about the mean of the homozygous gentotypes, as shown in Table 2. Covarying half the sums with half the differences of pairs gives cov $\frac{1}{4}(t_1 + t_2)(t_1 - t_2) = \frac{1}{2}dgd_1 + \frac{1}{4}hgh_1$ assuming that t_1 is always greater than t_2, that is, $e_1 > gd_1$, or cov $\frac{1}{4}(t_1 + t_2)(t_1 - t_2) = \frac{1}{4}hgh_1$ if $t_2 > t_1$, that is, $e_1 < gd_1$. If $e_1 > gd_1$, then the covariance, and hence the correlation, will be zero only if $gd_1 = gh_1 = 0$, or $\frac{1}{2}dgd_1 = -\frac{1}{4}hgh_1$, while if $e_1 < gd_1$ they will be zero only if $gh_1 = 0$. Hence the existence of a covariance is a useful guide to the presence of gd_1 and gh_1 interactions between genotype and environments.

It is possible to complicate this simple model in a wide variety of ways without substantially affecting the above conclusion.

In terms of MAVA the foregoing test may be written as follows:

$$\text{Expectation } \tfrac{1}{2}(t_1 - t_2) = |we|$$

$$\text{Expectation } \tfrac{1}{2}(t_1 + t_2) = bh + wh + be$$

where

bh = between-family hereditary deviation from the mean,
wh = within-family hereditary deviation from the mean,
be = between-family environmental deviation from the mean,
we = within-family environmental deviation from the mean, and
$|we|$ means the size of we, irrespective of its sign.

Notice that in MAVA no attempt is made to specify gene action effects such as d and h.

TABLE 2

TWELVE TWIN PHENOTYPES ARISING FROM THE INTERACTION OF GENOTYPE AND ENVIRONMENT

Twin frequency	Family	Phenotypes t_1	Phenotypes t_2	½ Sum $= \frac{1}{2}(t_1 + t_2)$	½ Differences $= \frac{1}{2}(t_1 - t_2)$
¼AA	1	$d + e_1 + gd_1 + e_2 + gd_2$	$d - e_1 - gd_1 + e_2 + gd_2$	$d + e_2 + gd_2$	$e_1 + gd_1$
	2	$d + e_1 + gd_1 - e_2 - gd_2$	$d - e_1 - gd_1 - e_2 - gd_2$	$d - e_2 - gd_2$	$e_1 + gd_1$
½Aa	3	$h + e_1 + gh_1 + e_2 + gh_2$	$h - e_1 - gh_1 + e_2 + gh_2$	$h + e_2 + gh_2$	$e_1 + gh_1$
	4	$h + e_1 + gh_1 - e_2 - gh_2$	$h - e_1 - gh_1 - e_2 - gh_2$	$h - e_2 - gh_2$	$e_1 + gh_1$
¼aa	5	$-d + e_1 - gd_1 + e_2 - gd_2$	$-d - e_1 + gd_1 + e_2 - gd_2$	$-d + e_2 - gd_2$	$e_1 - gd_1$
	6	$-d + e_1 - gd_1 - e_2 + gd_2$	$-d - e_1 + gd_1 - e_2 + gd_2$	$-d - e_2 + gd_2$	$e_1 - gd_1$
Mean		$\frac{1}{2}h + e_1 + \frac{1}{2}gh_1$	$\frac{1}{2}h - e_1 - \frac{1}{2}gh_1$	$\frac{1}{2}h$	$e_1 + \frac{1}{2}gh_1$

Thus on the one gene simple biometrical model

$$\text{Expectation covariance } \tfrac{1}{2}(t_1 - t_2)\,\tfrac{1}{2}(t_1 + t_2) = \text{cov}\,(bh + wh)\,|we| + \text{cov}\,(be)\,|we| = \tfrac{1}{2}dgd_1 + \tfrac{1}{4}hgh_1 + \text{cov}\,(e_2)\,|e_1|.$$

Written in this way, cov $(e_2)\,|e_1|$, which represents the interaction of family environmental influences with those within the family, is seen to be confounded with any genotype-environment interaction present. It will be observed that in the previous treatment this effect has been assumed absent.

When the monozygotic twins have been reared apart

$$\text{Expectation covariance } \tfrac{1}{2}(t_1 - t_2)\,\tfrac{1}{2}(t_1 + t_2) = \text{cov}\,(bh + wh)\,|we + be|$$

which is free of cov $(be)\,|we|$ bias. The rationale underlying this test of genotype-environment interaction is precisely that which underlies the test involving inbred lines of animals, and which detects genotype-environment interaction through heterogeneity of within-strain variances. In this case we are looking at heterogeneity of within-twin standard deviations caused by mean and standard deviation being related. By plotting scatter diagrams of $\frac{1}{2}(t_1 - t_2)$ and $\frac{1}{2}(t_1 + t_2)$, curvilinear covariance may be detected and subsequently tested for in addition to the simple linear covariance, thus extending the scope of the test to cover many possible forms of interaction.

Tests for Correlated Environments

If correlated environments exist, it can be shown, following Loehlin's (1965) method for deriving expectations of σ^2 MAVA, that the following expectations hold for σ_T^2 defined as $\sigma_W^2 + \sigma_B^2$ *in the analysis of variance* (which it will be remembered is $\frac{1}{2}\sigma_T^2$ MAVA).

For biological families raised together

$$\sigma_T^2 = \sigma_{wh}^2 + \sigma_{bh}^2 + \sigma_{we}^2 + \sigma_{be}^2 + 2r_{wh,we}\sigma_{wh}\sigma_{we} + 2r_{bh,be}\sigma_{bh}\sigma_{be}$$

For biological families raised apart (both sibs separated)

$$\sigma_T^2 = \sigma_{wh}^2 + \sigma_{bh}^2 + \sigma_{we}^2 + \sigma_{be}^2 + 2r_{wh,we}\sigma_{wh}\sigma_{we} + 2r_{bh,we}\sigma_{bh}\sigma_{we}$$

where σ_{wh}^2, σ_{we}^2 etc. are defined according to Cattell (1960). Now except under exceptional circumstances of internal balancing, these two σ_T^2s will not be expected to be equal unless the correlations contribute only an insignificant amount of covariance to the respective total variances.

Consequently, an F test for pooled σ_T^2 apart, against pooled σ_T^2 together, may be expected to indicate the importance of correlated environments. The above expectations can be complicated further by introducing placement correlations, special twin environments, and other types of family grouping without appreciably affecting the above argument. Where a number of σ_T^2s are to be tested, conventional tests for heterogeneity of variance may be employed (Winer, 1962). A further test can be made at the stage of fitting the model which is described in the next section.

It should be pointed out these two tests for the two kinds of f(G,E) are independent of each other, for the contribution of σ_{GE}^2 to σ_T^2s is the same for all kinds of families, and so cannot lead to heterogeneous total variances.

Fitting Models Assuming No f(G,E)

Models and Assumptions

In the absence of genotype-environment interactions the total variation can be partitioned between two components, the genetic G, and the environmental E. However, when we partition the total variation into σ_W^2 and σ_B^2 we must also paritition the total genetic and total environmental parts into within- and between-family portions. These we will designate:

G_1 = within-family genetic component,
G_2 = between-family genetic component,
E_1 = within-family environmental component,
E_2 = between-family environmental component.

These components have their direct equivalents in the four MAVA components σ_{wh}^2, σ_{bh}^2, σ_{we}^2, and σ_{be}^2, respectively. The expectations for the σ_W^2s and σ_B^2s for three different kinds of families are summarized in Table 3. The cumulative

assumptions embodied in the expectations are:

1. No genotype-environment interaction,
2. No correlated environments,
3. $G = G_1 + G_2$, and
4. E_1s and E_2s are the same for all kinds of families.

Assumption 1 can be tested before the model is fitted for twins reared together and apart by the methods described in the previous section (see *Tests for Genotype-Environment Interaction*). Assumptions 2, 3, and 4 can be tested, simultaneously, before the model is fitted by the heterogeneity, or otherwise of the σ_T^2s, by an analysis of variance of the means for each kind of family, and during the fitting of the model by a "goodness-of-fit" test of the estimated parameters.

To estimate the four parameters G_1, G_2, E_1, and E_2, and to test the four assumptions, a minimum set of data required comprises monozygotic twins reared together (MZ_T) and apart (MZ_A), and dizygotic twins reared together (DZ_T). This set provides six observed σ^2s which are sufficient to estimate four parameters and leaves two degrees of freedom for testing the equality of the three total σ^2s.

This set is not, however, the only minimum set of statistics which will lead to a solution, a point we will return to in a future section of this review (see *Minimum Data*).

TABLE 3

EXPECTATIONS OF VARIANCE COMPONENTS FOR THREE KINDS OF FAMILIES ACCORDING TO A SIMPLE GENETICAL MODEL

Monozygotic twins reared together (MZ_T)	
$\sigma_W^2 = E_1$	
$\sigma_B^2 = G + E_2$	
$\sigma_T^2 = G + E_1 + E_2$	
Monozygotic twins reared apart (MZ_A)	
$\sigma_W^2 = E_1 + E_2$	
$\sigma_B^2 = G$	
$\sigma_T^2 = G + E_1 + E_2$	
Dizygotic twins reared together (DZ_T) or full sibs reared together (FS_T)	
$\sigma_W^2 = G_1 + E_1$	
$\sigma_B^2 = G_2 + E_2$	
$\sigma_T^2 = G_1 + G_2 + E_1 + E_2$	

Heritability

Fitting the model and obtaining estimates of the parameters, however, is only the first step in the interpretation. We must consider the relationship of these parameters to heritability, and to the nature of the gene action involved. The only relevant heritability that can be obtained from the type of model fitting described so far is the so-called broad heritability of quantitative genetics which is $G/(G + E_1 + E_2)$, or $(G_1 + G_2)/(G + E_1 + E_2)$, that is, the proportion of the total variation due to all genetic causes. This ratio has no equivalent in either the classical or the MAVA approaches. Nevertheless, the so-called heritabilities of the classical approach can be expressed in terms of the parameters of the biometrical model. For example,

Holzinger's (1929):

$$H = (r_{MZ} - r_{DZ})/(1 - r_{DZ}) = G_1/(G_1 + E_1)$$

Nichol's (1965):

$$HR = 2(r_{MZ} - r_{DZ})/r_{MZ} = 2G_1/(G + E_2) = 2G_1/(G_1 + G_2 + E_2)$$

Vandenberg's (1966):

$$F = 1/(1 - H) = (G_1 + E_1)/E_1$$

These explicitly ignore important sources of variation (H ignores G_2 and E_2, for example). Similarly, the nature:nurture ratios of Cattell are equivalent to $\sigma_{wh}^2/\sigma_{we}^2 = G_1/E_1$ and $\sigma_{bh}^2/\sigma_{be}^2 = G_2/E_2$. Thus, if we can obtain estimates of G_1, G_2, E_1, and E_2 or their MAVA equivalents, we can not only estimate the broad heritability, but also the conventional heritabilities H, HR, etc., derived from the classical approach. Furthermore, if we can estimate only G, E_1, and E_2 we can still estimate the most useful heritability, broad heritability, even though we can no longer obtain the less useful H, HR, etc.

The Minimum Data

It is, therefore, important to establish the minimal experimental conditions under which we can estimate G_1, G_2, E_1, and E_2, or G, E_1, and E_2. These can be established by examining the expectations in Table 3.

MZ_A provide two statistics and require three parameters for their specification on the model. However, two of the parameters, E_1 and E_2, cannot be separated. Hence, we can estimate only G and $(E_1 + E_2)$. This, however, is sufficient to estimate the broad heritability alone and none of the other heritabilities, but in doing so it provides no test of the model. In the absence of genotype-environment interactions, MZ_T and MZ_A provide four statistics whose expectations can be expressed in terms of three parameters G, E_1, and E_2. We can, therefore, not only obtain least-squares estimates of these three parameters, but also have one degree of freedom in hand which can be used to test the equality of the total σ^2s for twins reared together and twins reared apart. This procedure effectively tests the adequacy of the model as well as providing the estimates.

If instead we have MZ_T and DZ_T we again have four statistics, but their expectations on our model involve all four parameters. Furthermore, two of the parameters, G_2 and E_2, occur only together in the expectations with the same coefficients, and are therefore inseparable. We can therefore estimate only G_1, E_1, and $(G_2 + E_2)$, leaving one degree of freedom for testing the equality of the total σ^2s for the two kinds of twins. We cannot estimate the broad heritability, but we can estimate, as has long been established empirically, *H*, *HR*, *F*, and one of the two nature: nurture ratios of Cattell.

If we have the combination of MZ_A and DZ_T we again have four statistics and four parameters, but we can obtain no joint solution of the parameters. Indeed, all we can obtain is the broad heritability from the monozygotic data alone, that is, the dizygotic twins can add nothing to the solution.

If we have all three sets of data, dizygotic and monozygotic twins, the latter reared together and apart, we have six statistics and four parameters, and all four can be estimated by least squares. The remaining two degrees of freedom allow us to test the equality of the three total σ^2s, and hence the validity of the model. With such data we are therefore in the position to estimate the broad heritability, *H*, *HR*, *F*, and both nature: nurture ratios. We can, of course, indefinitely extend the approach in this way to include more families of any kind, but these are in excess of the minimal requirements for a solution. We shall return to this point later.

A set of data which yields almost as much information as the above minimal set, and which is certainly far easier to collect, is comprised of MZ_T, DZ_T, and DZ twins (or sibs) reared apart (DZ_A or FS_A), the expectation for this latter group being

$$\sigma_W^2 = G_1 + E_1 + E_2$$
$$\sigma_B^2 = G_2$$
$$\sigma_T^2 = G_1 + G_2 + E_1 + E_2$$

This set yields estimates of G_1, G_2, E_1, and E_2 and allows tests of Assumptions 2, 3, and 4. The test for genotype-environment interaction (Assumption 1) is, however, incomplete, there being a test for gd_1 but not for gd_2. However, in view of the comparative ease of collecting such a sample, this cannot be considered a serious flaw. Moreover, the increased precision allowed by the larger sample size usually available makes this minimal set in some ways more useful than that which includes MZ_A. It may be objected that the use of sibs with monozygotic twins is not legitimate since it introduces an extra source of variation, that due to the inevitable age difference between sibs not present between the twins. This source of variance, if it exists, will be a further contribution to the heterogeneity of σ_T^2s, and is therefore tested for adequately during the procedures outlined above. Analysis of variance could also be used to detect this effect and an analysis of covariance, with age as a covariate, could be carried out prior to estimating the components of variance.

Thus by reducing the total variation into its within- and between-component parts, and specifying their expected values in terms of genetic and environmental components, we achieve two things: (*a*) We can see unambiguously how and why we can or cannot obtain various heritabilities from the data available and their relationships; and (*b*) we can not only achieve a biometrical genetical solution of the data, but encompass en route every other solution that has been proposed such as *H*, *HR*, and *F*, as given in the previous section (see *Heritability*.)

Gene Action

A further stage after converting the estimates of G and E into heritabilities, etc., is to attempt to relate them to the kind of gene action involved. This is the point at which biometrical genetics leaves all the other approaches including MAVA behind, because the starting point of a biometrical model is the nature of gene action and interaction. However, this is also the point at which the expectations depend on important assumptions about the kinds of gene action to be allowed for in the models and about the mating structure of human populations.

If we start with the simplest of all possible models, namely, one that assumes that all mating is at random (R), and all gene action is additive (d), the total genetic variation $G = \frac{1}{2}D_R$. In this expression $D_R = \Sigma\ 4u_iv_id^2$, where u_i is the frequency of the increasing allele at the ith locus, and v_i is the frequency of the decreasing allele at the same locus, and $u_i + v_i = 1$. The summation Σ is over all genes which are contributing to the variation of the character. In such a population $G = G_1 + G_2 = \frac{1}{2}D_R$, and $G_1 = G_2 = \frac{1}{4}D_R$. The broad heritability in such a population is $\frac{1}{2}D_R/(\frac{1}{2}D_R + E_1 + E_2)$, and the broad heritability equals the narrow heritability because we have excluded nonadditive gene effects.

Furthermore,

Holzinger's	$H = \frac{1}{4}D_R/(\frac{1}{4}D_R + E_1)$
Nichol's	$HR = \frac{1}{2}D_R/(\frac{1}{2}D_R + E_2)$
Vandenberg's	$F = (\frac{1}{4}D_R + E_1)/E_1$

and Cattell's nature:nurture ratios for within and between families are $E_1/\frac{1}{4}D_R$ and $E_2/\frac{1}{4}D_R$, respectively. None of these corresponds to the broad heritability of the population.

If we now allow for dominance effects (h) of the genes, the following changes must be made in the above expectations:

$$G = \tfrac{1}{2}D_R + \tfrac{1}{4}H_R$$
$$= \Sigma\ 2u_iv_i[d_i + (v_i - u_i)h_i]^2 + \Sigma\ 4u_i^2v_i^2h_i^2$$

Again, $G = G_1 + G_2$ but $G_1 \neq G_2$. In fact $G_1 = \frac{1}{4}D_R + \frac{3}{16}H_R$ and $G_2 = \frac{1}{4}D_R + \frac{1}{16}H_R$. Consequently, $G_1 - G_2 = \frac{1}{8}H_R$. The broad heritability $= (\frac{1}{2}D_R + \frac{1}{4}H_R)/(\frac{1}{2}D_R + \frac{1}{4}H_R + E_1 + E_2)$, which no longer equals the narrow heritability, this being $\frac{1}{2}D_R/(\frac{1}{2}D_R + \frac{1}{4}H_R + E_1 + E_2)$.

However we can still estimate the narrow heritability as well as the broad heritability if we have estimates of G_1, G_2, E_1, and E_2 and if the model holds because $\frac{1}{2}D_R = 3G_2 - G_1$. Then,

Holzinger's

$$H = (\tfrac{1}{4}D_R + \tfrac{3}{16}H_R)/(\tfrac{1}{4}D_R + \tfrac{3}{16}H_R + E_1),$$

Nichol's

$$HR = (\tfrac{1}{2}D_R + \tfrac{3}{8}H_R)/(\tfrac{1}{2}D_R + \tfrac{1}{4}H_R + E_2),$$

and Vandenberg's

$$F = (\tfrac{1}{4}D_R + \tfrac{3}{16}H_R + E_1)/E_1$$

which again do not correspond to any conventional heritability estimates, and, for reasons which will emerge later in this section, Cattell's nature:nurture ratios have no equivalents on this model using the method of estimation proposed by Cattell.

Although this model can be extended again to cover the complication of interactions between genes at different loci (see Fisher, 1918, and Kempthorne, 1957, for accounts of nonallelic interaction) these are unlikely to have such important consequences as other possible inadequacies of the model, such as the assumption of random mating. The most likely causes of deviation from random mating are (*a*) inbreeding, due either to a higher frequency of mating between relatives than expected under random mating, or to a higher frequency of mating than expected within small geographical areas of the population; and (*b*) positive assortative mating due to preferential mating of like phenotypes. To be effective, that is, to influence the genetical structure of the population, assortative mating must not only be preferential mating of like phenotypes, but also of like genotypes. Hence, if it is effective it leads to inbreeding. While, therefore, it is usual to consider the consequences of inbreeding, that is, mating between relatives independently of the consequences of assortative mating, both can be considered in terms of the homozygosity to which they lead.

Where we can control the mating of the individuals in the population under investigation, as we can with most animals, a past history of

inbreeding or assortative mating creates no great problem. By randomly mating individuals from the population, estimates of G_1 and G_2 may be obtained and random mating expectations of G_1 and G_2 modified by including f, the inbreeding coefficient of Wright (1951) (see Dickinson & Jinks, 1956; Jinks & Broadhurst, 1965). However, where the mating structure of the population cannot be changed, as in the case of humans, alternative methods must be used. Probably the most satisfactory of these approaches is that developed by Fisher (1918). If we have a population in which phenotypic assortative mating is taking place for a trait controlled by many genes of small effect, then at equilibrium the total genetic variance G will equal $\frac{1}{2}D_R + \frac{1}{4}H_R + \frac{1}{2}\{A/(1-A)\}D_R$ that is, the variance which would result from a random mating population with the same gene frequencies, $\frac{1}{2}D_R + \frac{1}{4}H_R$ plus a fraction of D_R, $\frac{1}{2}\{A/(1-A)\}D_R$ which is produced by the assortative mating. The constant A which has been introduced refers to the correlation between additive deviations of spouses, and is a simple function of their observed phenotypic correlation μ, the marital correlation. The A is 0 under random mating, and $A = \mu$ as an upper limit. The G will again equal $G_1 + G_2$ and

$$G_1 = \tfrac{1}{4}D_R + \tfrac{3}{16}H_R$$
$$G_2 = \tfrac{1}{4}D_R + \tfrac{1}{16}H_R + \tfrac{1}{2}\{A/(1-A)\}D_R$$

With random mating $G_1 > G_2$ by an amount equal to $\frac{1}{8}H_R$. With assortative mating and no dominance $G_1 < G_2$ by an amount $\frac{1}{2}\{A/(1-A)\}D_R$. Thus a significant difference between G_1 and G_2 will unambiguously detect either dominant gene action or assortative mating, but $G_1 = G_2$ will not necessarily indicate their absence since the effects of assortative mating and dominance will not necessarily lead to a difference between G_1 and G_2 when both are present to the same extent. In the absence of independent evidence of either dominance or assortative mating, it is probably reasonable, however, to accept $G_1 = G_2$ as indicating predominantly additive gene action. Where independent evidence of assortative mating is available, for example, through an observed marital correlation μ, it becomes possible to estimate the level of dominance by first estimating A, and then substituting into the expressions for G_1 and G_2, giving two equations which can then be solved for D_R and H_R. A number of ways of estimating A from μ are possible, that described by Fisher (1918) and used by Burt and Howard (1956) being to put $A = 2r_{p.o}\mu/(1+\mu)$ where $r_{p.o}$ is the parent-offspring correlation. However, $2r_{p.o}/(1+\mu)$ is merely an expression for the narrow heritability in an assortatively mating population given by $[\frac{1}{2}D_R + \frac{1}{2}\{A/(1-A)\}D_R]/\sigma_T^2$, so that an estimate of this heritability from whatever source enables us to estimate A as $A =$ (Heritability)μ.

We illustrate, in the first example on IQ (see *Fitting the Model*),a very simple iterative procedure for estimating D_R, H_R, and $\frac{1}{2}\{A/(1-A)\}D_R$ using $A =$ (Heritability)μ to obtain A. This method, which uses only estimates of G_1, G_2 and does not rely on $r_{p.o}$, a statistic likely to be biased upwards by common parental and offspring environments, leads to estimates very close to those obtained by Burt and Howard (1956) using $r_{p.o}$. A further method using $r_{p.o}$ (or covariance p.o where σ^2s are being used) is also illustrated. This latter method has the advantage over that used by Burt and Howard in that if $r_{p.o}$ is biased, then a conservative estimate of H_R is obtained and sensitive tests of significance are available for H_R, D_R, and $\frac{1}{2}\{A/(1-A)\}D_R$, whereas with other methods no tests of significance are possible. A similar approach to that of Fisher was developed by Wright (1921). Unfortunately, Wright did not allow for the presence of dominant gene action in his model, and so it must be deemed inferior to that of Fisher. Jensen (1967) has recently proposed a method allowing for the effects of assortative mating which involves the assumption that a previously random mating population has, in a single generation, had imposed upon it a degree of assortative mating. Under this model, which also ignores the effects of dominance, and appears to be restricted to a single gene effect, the correlation between siblings may rise from a possible maximum of 0.5 for a random mating population to 0.66 for an assortatively mating population, where the marital correlation $\mu = 1.0$, that is, where, with respect to a single gene, the mating system has changed from random mating to the extreme inbreeding situation of selfing. On this

model with human populations very small, effects are to be expected from assortative mating. On the other hand, with Fisher's model we may expect quite large effects, the sib correlation for $\mu = 1.0$ rising from a possible maximum of 0.5 to 1.0, a perfect correlation. In view of the highly restrictive assumptions implicit in Jensen's model, it may lead to quite misleading results where many genes are involved and assortative mating has been taking place for more than a single generation.

Having established the presence of dominant gene action, it remains to discover whether it is uni- or bidirectional in nature, this last step being perhaps the most important one in view of the implication for the evolutionary history of the trait. Traits which show strong directional dominance have probably been subject to strong directional selection during evolution, whereas those showing no dominance, or ambidirectional dominance have been selected for an intermediate optimum (Broadhurst & Jinks, 1966; Bruell, 1967; Fisher, Immer, & Tedin, 1932; Fulker, 1966; Mather, 1953; Roberts, 1967). Thus the current presence and direction of dominance indicates whether an intermediate or extreme level of expression of the trait has, in the past, been adaptively superior.

With controlled mating or using inbred lines it is a simple matter to determine the direction of dominance, but, unfortunately, with human populations there is a dearth of methods for measuring its direction. However, two would seem feasible. Perhaps the most direct method for detecting the direction of dominance is to examine the scores of children of consanguineous matings, for example, progeny resulting from cousin marriages. If we take a group of offsprings from nonconsanguineous matings, having first equated mean parental scores for this and the consanguineous group, then the mean of the offspring of the consanguineous group should show inbreeding depression if directional dominance is present.

The exact difference between the two groups is given by the equation

$$M_f = M_O - 2f\Sigma uvh$$

where

M_f = mean of inbred offspring,

M_O = mean of noninbred offspring,

and

f = Wright's inbreeding coefficient.

This formula would almost certainly apply very closely in the presence of a certain amount of assortative mating. We provide an example of this method in a subsequent section of this review. The second method is perhaps more easily applied, since it involves only scores for ordinary families of size three or more. Fisher, Immer, and Tedin (1932) showed that a number of the third moments of populations derived from two inbred lines, may, in the absence of genotype-environment interaction (which they refer to as "metrical bias") be used to detect the direction of dominance. Unfortunately, when the expectations of these third moments are derived for random mating populations, where unequal gene frequencies almost certainly exist, the effects of directional dominance and unequal gene frequencies cannot be disentangled. There appears, however, to be one exception, the mean skewness of within-family scores. Under random mating this has the expectation $\bar{k}_3 = -3\Sigma u^2v^2d^2h$. The $\bar{k}_3$ is negative where there is dominance for high expression of the trait, and positive where dominance is for low expression. The $\bar{k}_3$ is calculated from the mean of $\{n/(n-1)(n-2)\}\Sigma(x-\bar{x})^3$ for each family of size n, $(n > 2)$. Other possibilities involving third degree statistics may exist which we have not explored. The effect of assortative mating on this expectation is not known, but is expected to be small. For both these tests legitimate and sensitive tests of significance exist.

Further Statistics

In the types of families considered so far, namely, monozygotic and dizygotic twins, we have seen that our biometrical parameters G_1, G_2, E_1, and E_2 are directly relatable to the σ^2 MAVA for within- and between-family heredity and environment, respectively. However, this relationship breaks down if we consider other types of families because they are inadequately specified on the MAVA system. For example, whereas, assuming random mating and only additive and dominance effects of the genes, we can specify all the purely genetic contributions to the σ_W^2s and σ_B^2s of twins and full

siblings in terms of G, G_1, and G_2 where

$$G = \tfrac{1}{2}D_R + \tfrac{1}{4}H_R,$$

$$G_1 = \tfrac{1}{4}D_R + \tfrac{3}{16}H_R,$$

and

$$G_2 = \tfrac{1}{4}D_R + \tfrac{1}{16}H_R$$

we cannot specify the σ_W^2s and σ_B^2s for half-siblings. For them we require two new genetical components, G_3 and G_4. Even in this simple situation $G_3 = \tfrac{1}{4}D_R + \tfrac{1}{4}H_R$ and $G_4 = \tfrac{1}{4}D_R$. Indeed only when H_R is zero, that is, there are no dominance effects of the genes, can the genetic contribution to the variation within and between half-sib families be specified in terms of G_1 and G_2, and under these conditions $G_1 = G_2 = G_3 = G_4$. Thus Cattell's specification of half-sib families in terms of σ_{wh}^2 and σ_{bh}^2 implicitly assumes both random mating and genes with additive effects only. All the indications, both from the data analyzed later in this paper, and from the data analyzed by MAVA, are that the conditions under which the Gs are equal may not occur in practice. Should it prove necessary to extend the models to include other types of relationships, for example, cousins, parent-offspring, uncle-niece, etc., further types of Gs, which again are not equal under likely conditions, will have to be introduced. The value of these further statistics lies in the increasing power of the model to predict gene action since each G has an expectation in terms of D_R, H_R, and, if assortative mating is taking place, in terms of f or A as well.

Fitting Models When There is f(G,E)

Correlated Environments

Cattell (1960) and Loehlin (1965) have covered this topic at length so that we confine ourselves to a few comments. Although we accept that the kind of model MAVA proposes is in general appropriate, we suggest that the model fitting procedure be modified to accommodate a least-squares estimation procedure. This will help ensure that no superfluous terms are retained in the model. Nonsignificant items could be dropped from the model and the simplest adequate model fitted. In any case, it would seem (Loehlin, 1965) that $r_{whwe}\sigma_{wh}\sigma_{we}$ and σ_{wh}^2 are inseparable, so these must be fitted as a single compound term $(\sigma_{wh}^2 + 2r_{whwe}\sigma_{wh}\sigma_{we})$. The investigation of gene action would seem to be very difficult if significant correlations such as r_{whwe} exist.

Genotype-Environment Interaction

When f(G,E) is due to σ_{GE}^2 we have two courses open to us. The simplest is to rescale the data so as to minimize the interaction. In general the strength of the transformation will be indicated by the form of $r_{\frac{1}{2}\text{diff}\,\frac{1}{2}\text{sum}}$. For monozygotic twin groups, for example, with a strong linear relationship, scores may be transformed from x to $\sqrt[n]{x}$, where n = a number between 1 and 2. When the correlation detects a linear relationship between $\frac{1}{2}$sum and $\frac{1}{2}$diff^2, transformation of x to log x is indicated. However, although rescaling is possible, it may be more rewarding to pursue the analysis and interpretation of the interactions.

Since this requires the fitting of a model which allows for the effects of the interactions, this alternative requires data from many more types of relationships than does the fitting of the simpler model. With only MZ_T and MZ_A it is not possible to partition the interactive components from the environmental ones although the genetic component can still be estimated without bias, provided, of course, that separated twins are randomly distributed across environments.

Values of G, $E_1 + GE_1$, and $E_2 + GE_2$ may be obtained by the procedure adopted for estimating G, E_1, and E_2 in the absence of interaction, and the broad heritability may be precisely estimated by $G/(G + E_1 + GE_1 + E_2 + GE_2)$. It is perhaps worth noting that the argument that the possible presence of genotype-environment interaction invalidates the partitioning of variance approach to the nature: nurture problem is not entirely correct. The proportion of variance which is purely genetic in origin may be estimated, whether interaction is present or not.

To partition G into G_1 and G_2 in order to investigate gene action and the possibility of assortative mating, one further group (DZ_A) is necessary since the expectations for this group are $\sigma_W^2 = G_1 + E_1 + GE_1 + E_2 + GE_2$, and $\sigma_B^2 = G_2$. As in the expectations for MZ_A (see Table 3) the between-family component is unbiased by components of genotype-environ-

ment interaction when individuals are randomly distributed across all possible environments. From $\hat{G}_1$, $\hat{G}_2$, $\widehat{E_1 + GE_1}$, and $\widehat{E_2 + GE_2}$, appropriate forms of broad and narrow heritability may be obtained simply by replacing E_2 by $E_2 + GE_2$, and E_1 by $E_1 + GE_1$ in the previous formulas. The biometrical interpretations of these two heritability ratios as that portion of total variance which is genetic and that which is available for selection, respectively, still hold.

The complete separation of the six components in the model may not, however, be possible for it would appear that E_1 and GE_1 are inevitably confounded in the presence of family groupings (Mather & Morley-Jones, 1958).

The features of family structure which lead to the confounding of GE_1 with E_1 are precisely those which lead to the inevitable confounding of σ_{wh}^2 and $2r_{whwe}\sigma_{wh}\sigma_{we}$ that Loehlin (1965) had noticed in the MAVA equations. However we can separate E_2 and GE_2 by adding a further group to the MZ_T, MZ_A, and DZ_A, namely unrelated individuals reared together (U_T), since the expectations for this group are $\sigma_W^2 = G + E_1 + GE_1 + GE_2$, and $\sigma_B^2 = E_2$. In this case it is the random grouping of genotypes which allows an unbiased estimate of E_2. Although GE_1 cannot be estimated from the expectation of family variances, we can gain some insight into its likely size by using the variance of within-family deviations about the mean of such deviations, for MZ_T, as an upper bound for GE_1, and the product of this upper bound with the square of the coefficient of correlation used in the detection of GE_1 as setting a lower bound. These are only very rough indications, however, since the upper bound will be inflated by the sampling variance of e_1s and the lower bound will account for the linear portion only of the interaction.

The cumulative assumptions in the model extended to include genotype-environment interaction are

1. No correlated environments
2. $G_1 + G_2 = G$
3. The E_1s, E_2s are the same for all kinds of families.

A fourth assumption, that the GE_2s and GE_1s are the same for all types of families follows from Assumptions 2 and 3.

These assumptions can be tested simultaneously by all tests for the equality of σ_T^2s, and Assumptions 2 and 3 further by the analysis of variance of the means. It should, perhaps, again be emphasized that while correlated environments distort the σ_T^2s, GE_1 and GE_2 do not, the *total* interaction being the same for all kinds of families. A further investigation of genotype-environment interaction could be undertaken by partitioning GE_1 into G_1E_1 and G_2E_1, and GE_2 into G_1E_2 and G_2E_2.

However, only the components of GE_2 may be estimated. These components can be interpreted in terms of gene action (Mather & Morley-Jones, 1958), and would seem to mimic G_1 and G_2 in reflecting dominance and assortative mating. That is, in the presence of dominance $G_1E_2 > G_2E_2$, and in the presence of assortative mating the reverse is expected, $G_2E_2 > G_1E_2$. To estimate G_1, G_2, $(E_1 + G_1E_1 + G_2E_1)$, E_2, G_1E_2, and G_2E_2 the addition to the above four kinds of families of DZ_T becomes necessary.

The expectation for DZ_T is: $\sigma_W^2 = G + (E_1 + G_1E_1 + G_2E_1) + G_1E_2$, and $\sigma_B^2 = G_2 + E_2 + G_2E_2$. The terms inside the parentheses are inseparable. To estimate these six parameters it is only necessary to substitute G_1E_2 and G_2E_2 in the place of GE_2 in previously given expectations and to use least squares. There will be eight statistics and six parameters leaving 2 *df* for assessing the adequacy of the model.

One point perhaps worth mentioning in the estimation of genotype-environment interaction is that although it biases only environmental components, given that appropriate groups are chosen, in some cases it will bias genetic components if certain groups are used and undetected interaction is present. Estimation of G_1, G_2, E_1, and E_2 from MZ_T, MZ_A, and DZ_T will lead in fact to estimates of the following:

$$(G_1 + G_1E_2),$$

$$(G_2 - G_1E_2),$$

$$(E_1 + G_1E_1 + G_1E_2),$$

and

$$(E_2 + G_1E_2 + G_2E_2),$$

respectively. Thus although broad heritability is still unbiased, because $G = (G_1 + G_1E_2)$

$+ (G_2 - G_1E_2) = G_1 + G_2$, assessment of gene action becomes uncertain.

Both Correlated Environments and Genotype-Environment Interaction

Certainly the simplest course open to the investigator faced with the problem of both correlated environments and genotype-environment interaction is to rescale his data in an attempt to reduce the interaction to an insignificant level. Then if correlated environments still exist (and, in general, rescaling would not be expected to remove their effects), MAVA expectations can be fitted in the usual manner.

If a full model is required including GEs and $r_{h.e}$s then the expectations become very complex. We not only have the MAVA and interactive components to add to the model, but extra correlations to account for the correlations of the genetic *and* environmental deviations with the interactive ones. However, some simplifications may be possible by assessing the relative importance of the various components before fitting the model and dropping terms of negligible importance. Moreover, since within-family heredity correlations and E_1 and GE_1 are inseparable, some kind of simplified model has to be fitted in any case. Until it is clear that such complications as substantially correlated environments and considerable genotype-environment interaction exist simultaneously, it would not seem worth formulating the expectations. Cattell (1963) has proposed a method for introducing a scale factor k into the MAVA expectations to allow for genotype-environment interaction effects, a device which might prove useful for amending the biometrical genetical expectations as well. It would appear, however, that this form of correction is, in principle, little different from rescaling the original data before entering the analysis, a procedure which might well lead to a much simpler form of analysis. Before leaving the subject of f(G,E) it would seem worth making a few general points concerning its importance.

We have seen that if GE exists, it tends to bias components in such a way as to make corrections to heritability formulas automatic. Thus for the purposes of predicting the results of selection, either artificial or natural, the possible presence of GE is unlikely to lead to incorrect predictions. A similar argument would seem to hold for correlated environments. When they exist, the confounding of the covariance appears to be with the genotypic components so that for such predictors of population dynamics as heritabilities the correct answer is again obtained.

What would seem to be crucial in deciding whether or not to separate correlated environment covariance from the genotypic variance is whether or not the correlated effects are likely to be separable in practice. Cattell (1963) has argued persuasively for their importance, and examined with considerable ingenuity how they could arise. However, it is still not clear to us that many of the kinds of processes that he describes are meaningfully separated from the direct effects of genotype. To give one brief example: An innately intelligent person may well select his environment so as to produce positive r_{whwe}, and likewise a dull person may produce the same correlation by selecting less stimulating features of his environment. But is not this a more or less inevitable result of genotype? To what extent could we ever get a dull person to select for himself an intellectually stimulating environment to the same extent as a bright person might? Even when these correlations exist because of the pressure of others on an individual, it is not clear to what extent the correlation can be manipulated. Perhaps it can to some extent by such drastic procedures as intensively coaching the dull, and drastically depriving the intelligent, but the effect on the correlation is still not entirely clear. We need much more evidence concerning the effects and causes of correlated environments in order to decide their importance. In the meantime it might prove more realistic to adopt a "black box" approach, as suggested by Roberts (1967), and to consider all genotype correlated effects as truly genotypic and the residual effects as environmental, especially as regards effects operating within the family.

The concept of the modifiability of genotypes leads to an important distinction between the two sources of f(G,E). Both correlated environments and genotype-environment interaction modify genotypes and alter their relative differences, but they achieve this in quite different ways. The presence of correlated environments means that the relative differences between

genotypes have been altered by supplying to them, or perhaps more importantly, allowing them to accrue to themselves, precisely those environmental encounters needed to produce the relative differences observed in the phenotype. Each genotype has had a unique set of environmental encounters. In the case of negative correlation of environments, genetic differences are in a sense self-correcting so that individuals tend to become alike, a process Cattell has referred to as "coercion to biosocial norm" (1963). Where, as in the case of intelligence-test scores, the correlation seems likely to be positive, genotypic differences are accentuated. In either case, each individual genotype gets a unique environmental "treatment." As we have suggested, the implications of this process for devising methods to manipulate genotypes are not clear.

The presence of genotype-environment interaction, on the other hand, indicates a much simpler process in which the relative differences between genotypes are altered, not by providing each with a unique environment, but by supplying all with one of a number of possible uniform environments. By changing the regimen for all, the relative differences between all genotypes will be altered. This simpler process clearly has implications for social engineering. In the presence of correlated environments, environmental encounters would have to be redistributed, each according to the genotype's requirements, in order to effect change. This may not only be impossible, but, even if possible, quite unacceptable socially. In the presence of genotype-environment interactions, one particular set of environmental encounters, uniformly applied, may achieve the required change in relative differences. Moreover, if the genotype-environment interaction is detected by the correlation method previously described, the direction of change is also indicated. Unfortunately, an apparent lack of evidence of substantial genotype-environment interaction in intelligence-test scores strongly suggests that none of the range of environments provided by our society is likely uniformly to produce a high (or low) level of intelligence. The importance of trying to detect genotype-environment interaction in different societies, as a means of assessing their relative efficacy in achieving this end, is clearly indicated.

The Classical Approach Through Correlations

The intraclass correlations used in the classical approach (Burt, 1966; Fisher, 1918; Huntley, 1966; Husén, 1959; Newman, Freeman, & Holzinger, 1937), are formally equivalent to variance components with the restriction that each of the σ_T^2s of the groups used has been brought to a common base.

We have seen that $\hat{r} = \hat{\sigma}_B^2/(\hat{\sigma}_B^2 + \hat{\sigma}_W^2)$ or $\hat{\sigma}_B^2/\hat{\sigma}_T^2$, where the symbol $\hat{r}$ refers to the estimate of r and $\hat{\sigma}_B^2$ to the estimate of σ_B^2, etc., r is, therefore, simply σ_B^2 expressed as a fraction of the total variance for that group. Thus we are able to fit genetical models to rs in the same way that we are able to fit them to σ_B^2 and σ_W^2.

For example:

	G	E_1	E_2
MZ_T, r =	1	0	1
MZ_A, r =	1	0	0

E_1 being obtained as $1 - G - E_2$. The G, E_1, and E_2 have unique solutions because the correlation brings all statistics to a common base of 1, and so a least-squares solution is not required. The G_1, G_2, E_1, and E_2 may be found by adding DZ_T.

	G_1	G_2	E_1	E_2
MZ_T, r =	1	1	0	1
MZ_A, r =	1	1	0	0
DZ_T, r =	0	1	0	1

Again, $E_1 = 1 - G_1 - G_2 - E_2$, and again a least-squares estimation is not required. Further groups could be added to detect further parameters, G_3, G_4, etc., as previously specified and the detection of gene action and assortative mating attempted. Also, GE parameters may be fitted by adding further groups. In fact, any parameters which do not lead to inequality of the σ_T^2s may be allowed into the model.

Using $\hat{r}$ instead of σ^2s assumes equality of σ_T^2s, so that all failures of the model which lead to their inequality must be assumed absent, given adequate sampling. These, as we have seen, are

1. $G_1 + G_2 \neq G$
2. E_1s and E_2s not equal for all groups
3. Correlated environments

Failure to test for the presence of these effects

in their data by those who use this approach—although, of course, they could do so by preliminary tests for heterogeneous variances—will prospectively bias all the parameters in a highly complicated way depending on any relative inequalities of σ_T^2s present. Rather than attempt to bring this correlational approach in line with biometrical genetics and MAVA it would seem simpler to abandon it and work in terms of the variance components σ_W^2 and σ_B^2 directly. However, as we shall see, it can be useful for the purpose of reanalysis of published correlations provided, we bear in mind its limitations.

Examples—Reanalysis

The eight phenotypes chosen to illustrate the procedures outlined in the previous section

TABLE 4

Description of Data Chosen for Illustrative Reanalysis

Number	Phenotype	Description of test used	Data available	Source
1	Neuroticism	Self-rating questionnaire, constructed for Shields' study and similar to the Maudsley Personality Inventory	MZ_T (29 F and 14 M pairs of subjects) MZ_A (26 F and 14 M pairs of subjects) DZ_T (16 F pairs of subjects)	Shields (1962)
2	Extraversion	Same as above	Same as above	Shields (1962)
3	Intelligence	Synonyms section (Set A) of Mill Hill Vocabulary Test (Form B, 1948)	MZ_T (24 F and 12 M pairs of subjects) MZ_A (25 F and 15 M pairs of subjects)	Shields (1962)
4	Intelligence	Dominoes Intelligence Test. Nonverbal 20-minute test similar to Raven's progressive matrices	MZ_T (23 F and 11 M pairs of subjects) MZ_A (24 F and 14 M pairs of subjects)	Shields (1962)
5	Intelligence quotient	Group test standardized by the London Revision of Terman-Binet Intelligence Scale	Correlations only for: MZ_T (95 pairs of subjects) MZ_A (53 pairs of subjects) DZ_T (127 pairs of subjects) FS_T (264 pairs of subjects) FS_A (151 pairs of subjects) U_T (136 pairs of subjects) *Parents and offspring* (numbers unknown) Husband and wife (numbers unknown)	Burt (1966); Burt & Howard (1956)
6	Educational attainments	Group test devised for use with London school children, including reading, spelling, and arithmetic items. (Burt, 1921)	As above, (excluding last two correlations)	Burt (1966)
7	Intelligence quotient	IQ Scores deriving from a number of tests, preference given to individual tests administered at age 14	Sibs together (689 families; size > 3)	Reed & Reed (1965)
8	Intelligence quotient	Japanese version of the Wechsler Intelligence Scale for Children	1511 inbred children ($f = \frac{1}{16}$); 1608 control children ($f = 0$)	Spuhler (1967)
9	Intelligence quotient	Stanford-Binet Intelligence Scale and Otis Group Intelligence Scale	MZ_A (19 pairs of subjects; both sexes)	Newman, Freeman, & Holzinger (1937)

Note.—Abbreviations are: M = male; F = female.

of this paper are listed in Table 4. None of these sets of data is wholly satisfactory to illustrate our point of view. In some cases critical groups were nonexistent, or numbers in them small, and in others the raw scores were not available. However, the examples given cover a wide range of the practical problems of analysis and interpretation, in spite of their often severe limitations.

Neuroticism

The first trait chosen is neuroticism (Shields, 1962). The trait was measured by means of a 38-item self-rating questionnaire designed to give a measure of both neuroticism and extraversion. The questionnaire was specially constructed for the study by H. J. Eysenck, and is apparently similar to the Maudsley Personality Inventory (MPI). In so far as the trait in this example resembles neuroticism as measured by the MPI, it refers to a general emotional instability with a tendency to neurotic breakdown under stress, and is the name Eysenck gives to a broad, second order factor which, together with an independent factor labeled Extraversion, accounts for most of individual differences in the personality domain (Eysenck, 1960b). Unfortunately, however, it is not known how closely Shield's (1962) test and the MPI resemble each other.

Data from the following pairs of twins were available: 43 pairs of MZ_T, 40 pairs of MZ_A, and 16 pairs of DZ_T. Unfortunately the groups of males were very small, and only the female data approached a satisfactory volume for adequate analysis. In general, we have included the male data, where they agree with the female, in order to provide replication and augmentation of them. Where there was serious disagreement between male and female data, the males were discarded. In consequence, our conclusions apply more reliably to females than to males. However, as may be seen, the agreement between sexes was satisfactory on the whole.

TABLE 5

ESTIMATES OF THE VARIANCE COMPONENTS FOR THREE KINDS OF FAMILIES (NEUROTICISM)

Type of family	Variance component	Estimate	
		Female	Male
MZ_T	$V_{\bar{F}}$	11.0819	—
	$\hat{V}_F$	8.1207	—
MZ_A	$V_{\bar{F}}$	14.5608	14.7307
	$\hat{V}_F$	9.6635	5.0000
DZ_T	$V_{\bar{F}}$	11.7828	—
	$\hat{V}_F$	13.8552	—

TABLE 6

TABLE OF MEANS AND VARIANCES FOR THE FOUR KINDS OF FAMILIES (NEUROTICISM)

Type of family	Mean	Variance
MZ_T		
female	9.7241	15.0191
MZ_A		
female	11.8558	19.2018
male	10.7143	16.7778
DZ_T		
female	10.3282	17.8002

Analysis of variance and estimates of variance components. The first step is an analysis of variance to obtain the within- and between-family variances. For the female MZ_T we have 29 pairs, that is, 58 individuals giving 57 *df* of which 28 are for variation between families and 29 for the variation within families. This gives a between-family $MS = 22.1638$, with expectation $\sigma_W^2 + 2\sigma_B^2$, and a within-family $MS = 8.1207$, with expectation σ_W^2. From the preceding we can see that there is a significant between-family variance (F for 28/29, $df = 2.73$, $p = .01$) and

$$\hat{V}_{\bar{F}} = \tfrac{1}{2}\sigma_W^2 + \sigma_B^2 = 11.0819$$

$$\hat{V}_F = \sigma_W^2 = 8.1207$$

The corresponding analyses of variance for the other kinds of twins yield the estimates of $V_{\bar{F}}$ and $\hat{V}_F$ listed in Table 5. One group, the 14 pairs of male MZ_T, fails to yield a significant between-family MS, unlike the female data, and was omitted from subsequent analysis. For this group and this trait there would appear to have been inadequate sampling of between-family differences.

Testing the assumptions. If the individuals in the samples of the four types of families have been drawn at random from the same population, they should have the same means and variances. This is readily ascertained from the means and variances of the four samples listed in Table 6. An analysis of variance to compare

the four means is given in Table 7. This analysis shows that while there are differences between families, there are no overall differences between the types. The types can be approximately partitioned into the three orthogonal comparisons shown. We see there is no evidence of a sex difference in the one group which allows such a comparison. We can therefore regard the two sexes in the MZ_A group as providing two replicate samples from this group, which in turn allows us to assess the error of estimation in fitting the biometrical models. There is some suggestion that twins reared apart are slightly more neurotic than those reared together, but the significance level is borderline. There is no evidence that monozygotic twins differ from dizygotics.

Turning to the variances in Table 6, we find no evidence of significant differences, although the slight differences tend to parallel the differences in the means, a fact suggesting that what differences there are result from a slight curtailment of the distribution rather than from any other causes. Overall, therefore, there is no compelling evidence for regarding the four kinds of families as samples from different populations.

The second assumption to test is the possible importance of f(G,E) in these data. To test for genotype-environment interaction we calculate the product-moment correlation between family sums and differences for the monozygotic twins. For MZ_T, $r_{41} = 0.1489$ which accounts for 2% of the variation and fails to reach significance. We conclude therefore that there is no evidence of GE_1. For MZ_A, $r_{38} = 0.0583$, suggesting that GE_2 is not present in these data either. To test for evidence of the correlated environments of MAVA, we must see if the variance of separated families differs from that when the twins were reared together.

The total mean variances are

All twins together $\sigma_T^2 = 13.8644$ for 117 *df*

All twins apart $\sigma_T^2 = 18.1303$ for 79 *df*

An F test gives $F_{79,117} = 1.31$, $p = 2 \times .1 = .2$, a nonsignificant value suggesting that correlated environments are not a complication present in these data. Fitting the model will yield a further test of the importance of this source of variation in the test for the goodness of fit.

TABLE 7

AN ANALYSIS OF VARIANCE TO COMPARE THE FOUR MEANS IN TABLE 6

Source	*df*	*MS*	*F*
Between types of families	3	50.0500	1.96
Between families within types	81	25.6026	2.81*
Within families	84	9.1235	
Between types of families partitioned			
MZ_A sex difference	1	42.6168	1.66
MZ_A versus MZ_T and DZ_T	1	95.5952	3.73
MZ versus DZ (F only)	1	11.9383	—

* $p < .001$.

Thus on the basis of these tests we are justified in fitting the simple G and E models to the data.

Fitting the model. To illustrate the method we will first fit the appropriate model to the monozygotic twins only. Because we are able to regard the two sexes as replicates we have two values of $\bar{V}_F$ and $V_{\bar{F}}$ for MZ_A. The model is to be fitted to the mean of these values. For MZ_T, where we have no replications, there is only one value each of $\bar{V}_F$ and $V_{\bar{F}}$ to which to fit the model.

Thus we have four equations and three unknowns. Since there are fewer unknowns than observed statistics, we can use a least-squares procedure to estimate G, E_1, and E_2. The normal equations are

$$\begin{bmatrix} 2 & 1 & 1\frac{1}{2} \\ 1 & 2\frac{1}{2} & 1\frac{3}{4} \\ 1\frac{1}{2} & 1\frac{3}{4} & 2\frac{1}{4} \end{bmatrix} \begin{bmatrix} \hat{G} \\ \hat{E}_1 \\ \hat{E}_2 \end{bmatrix} = \begin{bmatrix} 25.7277 \\ 28.3163 \\ 25.7365 \end{bmatrix}$$

By inversion of the matrix we obtain the following solution:

$$\begin{bmatrix} 1.025 & 0.150 & -0.800 \\ 0.150 & 0.900 & -0.800 \\ -0.800 & -0.800 & 1.600 \end{bmatrix} \begin{bmatrix} 25.7277 \\ 28.3163 \\ 25.7365 \end{bmatrix} = \begin{bmatrix} \hat{G} \\ \hat{E}_1 \\ \hat{E}_2 \end{bmatrix}$$

Whereupon

$$\hat{G} = 10.0291$$
$$\hat{E}_1 = 8.7546$$
$$\hat{E}_2 = -2.0568$$

From these figures we can calculate the expected values of the four statistics listed along with the observed values in Table 8. We have one degree of freedom for comparing the ob-

TABLE 8

OBSERVED AND EXPECTED VALUES OF VARIANCE COMPONENTS FOR MZ_T AND MZ_A (NEUROTICISM)

Type of family	Variance component	Model			Statistics				
					Observed			Expected deviation	
		G	E_1	E_2	Female	Male	Mean		
MZ_T	$\bar{V}_F$	0	1	0	8.1207	—	8.1207	8.7546	−0.6339
	$V_{\bar{F}}$	1	½	1	11.0819	—	11.0819	12.3496	−1.2677
MZ_A	$\bar{V}_F$	0	1	1	9.6635	5.0000	7.3318	6.6978	+0.6339
	$V_{\bar{F}}$	1	½	½	14.5608	14.7307	14.6458	13.3781	+1.2677

served and expected statistics, and this tests the equality of the total σ^2s for twins reared together and apart. From replication we have an error variance for 2 *df* against which to test the significance of the discrepancy between the two total σ^2s. Variance (observed-expected) = Sum of the squares of the deviations listed in Table 8. Thus, $V(O\text{-}E) = 4.0152$. The error sum of squares for the $\bar{V}_F$ and $V_{\bar{F}}$ is given by $\Sigma\frac{1}{2}$(sex difference)^2 = 10.8871 for 2 *df*, 5.4436 is the variance for two of the V_S, but $\frac{1}{2}$(5.4436) is the variance attaching to the two mean values representing twins reared apart. Thus the error for testing $V(O\text{-}E)$ is $\frac{1}{2}$ 5.4436 + $\frac{1}{2} \times \frac{1}{2}$ 5.4436 = 4.0827 for 2 *df*. Thus $F_{1,2} = V(O\text{-}E)/V$ (error) = 4.0152/4.0827 = 0.98, a nonsignificant value confirming the result of the earlier *F* test on the total variances for twins reared together and apart. A further advantage of replication is that we can derive standard errors for the estimates of G, E_1, and E_2. Because the variance of total σ^2s given by $V(O\text{-}E)$ is not significantly greater than the variance between $\bar{V}_F$s and $V_{\bar{F}}$s given by V(error), we are justified in pooling these two sources of variation to give a pooled error variance, V(error) pooled = 4.0602 for 3 *df*. Multiplying by the appropriate coefficient in the leading diagonal of the inverse of the coefficients of the normal equations given above, we obtain the error variance of each estimate in turn.

$$V(\hat{G}) = 1.025 \times 4.0602 = 4.1617$$
$$V(\hat{E}_1) = 0.900 \times 4.0602 = 3.6542$$
$$V(\hat{E}_2) = 1.600 \times 4.0602 = 6.4963$$

The estimated components of the model, their standard errors, and the significance of their difference from zero by *t* test are

$$\hat{G} = 10.0291 \pm 2.0400, \quad t_3\hat{G} = 4.92,\ p = .017$$
$$\hat{E}_1 = 8.7546 \pm 1.9116, \quad t_3\hat{E}_1 = 4.58,\ p = .020$$
$$\hat{E}_2 = -2.0568 \pm 2.5488, \quad t_3\hat{E}_2 = 0.81,\ p = .420$$

Thus despite lamentably inadequate replication, we can see that G and E_1 are clearly significantly greater than zero while the negative E_2 is not.

We can repeat the estimations using the combined statistics from the total sample of monozygotic and dizygotic twins. We now have six statistics and four parameters leading to the solution

$$\begin{bmatrix} 1.6 & -1.6 & -0.8 & 0.8 \\ -1.6 & 2.55\dot{} & 0.9 & -1.6 \\ -0.8 & 0.9 & 0.86\dot{} & -0.8 \\ 0.8 & -1.6 & -0.8 & 1.6 \end{bmatrix} \times \begin{bmatrix} 45.4743 \\ 37.5105 \\ 48.0629 \\ 37.5193 \end{bmatrix} = \begin{bmatrix} \hat{G}_1 \\ \hat{G}_2 \\ \hat{E}_1 \\ \hat{E}_2 \end{bmatrix}$$

so that

$$\hat{G}_1 = 4.3072$$
$$\hat{G}_2 = 6.1186$$
$$\hat{E}_1 = 9.0190$$
$$\hat{E}_2 = -2.0568$$

We can now compare the observed and expected values of the six statistics (see Table 9).

Sum of squares (observed-expected) = 6.1162 which now has 2 *df*; thus $V(O\text{-}E) = 3.0581$, while V(error) = 4.5363, slightly greater than before because it is now based on two mean values and four single observations. The F ratio for the equality of the three total $\sigma^2 = 3.0581/4.5363 = 0.67$, which again confirms the earlier impression of homogeneity of σ_T^2s.

Pooling sources of error as before gives V(error) pooled = 3.7972 for 4 *df*. Multiplying this value by the appropriate leading diagonal elements in the inverse matrix yields the following:

$$\hat{G}_1 = 4.3072 \pm 2.4649, \quad t_4\hat{G}_1 = 1.75,\ p = .16$$

$$\hat{G}_2 = 6.1186 \pm 3.1117, \quad t_4\hat{G}_2 = 1.97,\ p = .18$$

$$\hat{E}_1 = 9.0190 \pm 1.8141, \quad t_4\hat{E}_1 = 4.97,\ p = .001$$

$$\hat{E}_2 = -2.0568 \pm 2.4649, \quad t_4\hat{E}_2 = 0.83,\ p = .45$$

Thus we have significant E_1, but nonsignificant G_1, G_2, and E_2.

The nonsignificant E_2 agrees with the previous analysis, but the nonsignificant G_1 and G_2 do not. The disagreement, however, is more apparent than real. Thus, if the model is adequate, the highly significant $\hat{G}$ of the previous analysis should equal the sum of the nonsignificant $\hat{G}_1$ and $\hat{G}_2$ of the present analysis. In fact $\hat{G}_1 + \hat{G}_2 = 10.4258$, which is very close to the estimate of G which equals 10.0291 $\pm$ 2.0400.

The significance of $\hat{G}_1 + \hat{G}_2$ may be tested by means of our standard errors

$$\text{Variance } (\hat{G}_1 + \hat{G}_2) = V(\hat{G}_1) + V(\hat{G}_2) + 2 \text{ cov } (\hat{G}_1,\hat{G}_2)$$

The V(error) multiplied by the appropriate coefficients in the inverse matrix will give us $V(\hat{G}_1)$ and $V(\hat{G}_2)$ and the term cov $(\hat{G}_1,\hat{G}_2)$ is obtained by multiplying V(error) by the first off-diagonal term in the same matrix.

$$V(\hat{G}_1 + \hat{G}_2) = V(\text{error})\ \{1.6 + 2.55 - 3.2\} = 3.6073$$

$$\text{Standard error} = \sqrt{3.6073} = \pm 1.8988$$

TABLE 9

OBSERVED AND EXPECTED VALUES OF VARIANCE COMPONENTS FOR MZ_T, MZ_A, AND DZ_T (NEUROTICISM)

Type of family	Statistics	Mean observed	Expected	Deviation
MZ_T	$\bar{V}_F$	8.1207	9.0190	−0.8983
	$V_{\bar{F}}$	11.0819	12.8786	−1.7967
MZ_A	$\bar{V}_F$	7.3317	6.9623	+0.3695
	$V_{\bar{F}}$	14.6458	13.9070	+0.5458
DZ_T	$\bar{V}_F$	13.8552	13.3263	+0.5289
	$V_{\bar{F}}$	11.7828	10.7250	+1.0578

Thus the two estimates of G differ considerably less than their standard errors, hence there is no disagreement between the outcome of the analysis of the monozygotic data only and that of the combined data from monozygotic and dizygotic twins. The latter estimate, which is based on $\hat{G}_1 + \hat{G}_2$ is not only highly significant ($p < .001$), but its standard error is smaller than that of the alternative estimate of G, as might be expected, since more data is involved in the estimation.

Analysis of gene action and mating system. What implications do these estimates have for gene action and the mating system? The only indication on these two points to emerge from the analysis of components is that $G_2 > G_1$. According to the models described, this indicates assortative mating rather than any dominant gene action. However, before accepting this interpretation we test the significance of $\hat{G}_2 - \hat{G}_1$ by means of our standard errors:

$$\text{Variance } (\hat{G}_2 - \hat{G}_1) = V(\hat{G}_2) + V(\hat{G}_1) - 2 \text{ cov } (\hat{G}_1\hat{G}_2)$$

As before

$$V(\hat{G}_2 - \hat{G}_1) = V(\text{error})\ \{1.6 + 2.55 + 3.2\} = 27.9094$$

Therefore, standard error

$$= \sqrt{27.9094} = 5.2829, \text{ and}$$

$$\hat{G}_2 - \hat{G}_1 = 1.8114 \pm 5.2829, \quad t_4 = 0.30,\ p = .8$$

Thus, although assortative mating is indicated for this trait, it cannot be proved to be significant in these data. At the same time, the absence of dominant gene action is clearly

TABLE 10

VARIOUS HERITABILITY RATIOS AND INDEXES FOR NEUROTICISM

Index	Estimate	
	Monozygotic data only	Monozygotic and dizygotic data
Broad heritability	0.60 ± 0.11	0.54 ± 0.09
Narrow heritability	—	0.54 ± 0.09
Holzinger's *H*	—	0.37 ± 0.08
Nichol's *HR*	—	1.00
Vandenberg's *F*	—	1.58 ± 0.25
Cattell's nurture: nature		
overall	0.67 ± 0.30	0.84 ± 0.42
within family	—	1.68 ± 0.61
between family	—	0.00

Note.—The between-family nurture:nature ratio is zero because $E_2 = 0$ in the model, and Nichol's $HR = 1.00$ for the same reason. No error is appropriate to them under maximum likelihood estimation.

indicated by the fact that $G_1 \not> G_2$. The absence of dominant gene action strongly suggests that an intermediate level of neuroticism has been favored by natural selection, and constitutes the population optimum for this personality trait. Gottesman (1965) has speculated along these lines and suggested for a number of such traits that extremes would be at a selective disadvantage, but little by way of evidence has been previously available.

Computation of heritability ratios following simplified model. Before computing the heritabilities, it is worth considering how we may obtain more precise estimates of G_1, G_2, E_1, and E_2 from these data, given that $\hat{G}_1$ and $\hat{G}_2$ are not significantly different from each other, and E_2 is not significantly greater than zero. Thus we can fit a simplified model where $G_1 = G_2$ and $E_2 = 0$.

We now have six statistics and only two parameters leading to the solutions:

$$\begin{bmatrix} 0.1\dot{3} & -0.1\dot{3} \\ -0.1\dot{3} & 0.40 \end{bmatrix} \begin{bmatrix} 82.9848 \\ 48.0629 \end{bmatrix} = \begin{bmatrix} \hat{G}_{1,2} \\ \hat{E}_1 \end{bmatrix}$$

giving

$$\hat{G}_1 = \hat{G}_2 = 4.6562 \pm 0.8484,$$
$$t_6\hat{G}_{1,2} = 5.49,\ p < .002$$
$$\hat{E}_1 = 8.1605 \pm 1.4605,$$
$$t_6\hat{E}_1 = 5.59,\ p < .002$$

Our sum of squares for observed-expected now has 4 *df* to test the adequacy of this simplified model. Notice we are now not merely assessing the equality of the total σ^2s, but $G_1 = G_2$ and the equality of E_1s as well. Thus, $V(O\text{-}E) = 4.5964$ for 4 *df*, $F_{4,2}$ for adequacy of model $= 4.5964/6.8045 = 0.67$, which is clearly a nonsignificant discrepancy, and the simple model is judged adequate. This gives pooled error $V(\text{error}) = 5.3324$ for 6 *df* and standard errors as given above.

Probably the very best estimate of G_1, G_2, and E_1 may be obtained by a weighted least-squares procedure where each observed V is weighted by the amount of information we have about it. In practice we do not know precisely what the amount of information is, but if we use $1/V(V)$ we will obtain a good approximation to the ideal procedure, and obtain approximate maximum likelihood estimates of our Gs and Es. This is a technique based on a method due to Nelder (1960), and is explained and illustrated more fully in Example 5, Table 4.

The maximum likelihood method gives

$$\hat{G}_1 = \hat{G}_2 = 4.5845 \pm 1.2471,$$
$$c = 3.68,\ p < .001$$
$$\hat{E}_1 = 7.7199 \pm 1.2755,$$
$$c = 6.05,\ p < .001,$$

which agree very well with the simple method previously given. Each estimate may be tested against its standard error as a normal deviate (c) since these errors are theoretical values. Obviously we obtain a much more powerful test than by using unweighted least squares. The test of the fit of the model now leads to an approximate $\chi_4^2 = 1.3217$, $p = .8$, which again confirms the adequacy of the simple model. It is reassuring that the relatively simple unweighted procedure leads to the same conclusions and similar values for our estimates as the more laborious weighted procedure.

This simplified model, which fits the data extremely well, and yields highly significant estimates of $G_1 = G_2$ and E_1 with $E_2 = 0$, may now be used to calculate heritabilities with some degree of confidence (see Table 10). Standard errors for these estimates have been calculated by the method suggested by Kempthorne (1957) which employs the formula for

the variance of a ratio,

$$V(A/B) = (A/B)^2\{V(A)/A^2 - 2\,\mathrm{Cov}\,(AB)/AB + V(B)/B^2\}$$

upon simple rearrangement of Kempthorne's expression. Now this is the formula Cattell (1963, 1965) suggested using for calculating errors for nature:nurture ratios. Comparison of Cattell's formula with ours suggests a printing error in Cattell's 1965 paper. Cattell, however, would appear to be mistaken concerning the effect of the covariance term in this expression in claiming it makes $V(A/B)$ smaller than if Cov $(AB)/AB$ were ignored.

This covariance term is almost invariably negative (as may be seen from the inverse matrices of the normal equations given in our examples) and will therefore result in inflation of variance.

To illustrate the use of this expression we calculate the error for E_1/G_1, the within-family nature:nurture ratio. Here

$$V(\hat{E}_1) = 1.6269$$
$$V(\hat{G}_1) = 1.5553$$
$$\mathrm{Cov}\,\hat{E}_1,\hat{G}_1 = -0.5546,$$

these values being obtained from the inverse matrix involved in the estimation procedure.

Substituting in the expression above we obtain

$$V(E_1/G_1) = 1.68^2\{1.6269/7.7199^2 + 2(0.5546)/(7.7199)(4.5845) + 1.5553/4.5845^2\}$$

Therefore, $SE\ (E_1/G_1) = \sqrt{0.3742} = 0.6117$

Thus $E_1/G_1 = 1.68 \pm 0.61$ as given in Table 10. The striking feature of Table 10 is that all ratios are significant and capable, therefore, of interpretation.

The narrow heritability equaling the broad heritability all genetic variation (54% of the total variance) is available for natural or artificial selection to act upon. Cattell's nurture: nature ratios indicate that although environment is more important than genotype in producing differences between siblings, the differences in neuroticism observed between families is entirely genotypic in origin. Evidently cultural and class differences have no effect on this major personality dimension. For a discussion of the value of calculating other estimates, the reader is referred to the references under the appropriate authors.

TABLE 11

ESTIMATES OF VARIANCE COMPONENTS FOR THREE KINDS OF FAMILIES (EXTRAVERSION)

Type of family	Variance component	Estimate	
		Female	Male
MZ_T	$V_{\bar{F}}$	7.2384	10.6680
	$\bar{V}_F$	6.4871	6.5893
MZ_A	$V_{\bar{F}}$	15.3235	11.2857
	$\bar{V}_F$	7.6923	3.5714
DZ_T	$V_{\bar{F}}$	8.3123	—
	$\bar{V}_F$	26.5547	—

Extraversion

The next trait chosen for analysis is extraversion as measured by the self-rating questionnaire described in the previous example. Again, it is not clear how closely this trait resembles its counterpart in the MPI which refers to uninhibited, outgoing, and sociable tendencies in behavior but a moderate resemblance at least seems certain (Shields, 1962). This trait, together with neuroticism, completes the broad, two-dimensional view of major personality tendencies described by Eysenck (1960b).

Scores for the same individuals as in the previous example are available.

Analysis of variance and estimation of variance components. These calculations produced the components shown in Table 11.

Testing the assumptions. An analysis of variance to compare the means of the five types of families (see Table 12) is given in Table 13. This analysis shows that overall there is some suggestion of types differing, and that there are differences between families. The approximate orthogonal comparisons between types indicates that monozygotic twins reared apart are significantly less extravert than those reared together. There is no sex difference so we may again regard sex as replicates. There is no difference between monozygotic and dizygotic twins. Our two samples of monozygotic twins would seem to have been drawn from a different population with respect to their means. The

TABLE 12

TABLES OF MEANS AND VARIANCES FOR THE FIVE KINDS OF FAMILIES (EXTRAVERSION)

Type of family	Mean	Variance
MZ_T		
female	13.8017	10.4035
male	13.7321	13.6895
MZ_A		
female	11.7885	18.9445
male	11.3571	12.7195
DZ_T		
female	12.4844	21.7497

variance gives $S_{max}^2/S_{min}^2 = 2.09$, $p = .1$, which is only a borderline significance, while an F test for σ_T^2 tog versus σ_T^2 ap $= 1.18$, $p = .2$, strongly suggesting that correlated environments are not important in these data. The variances appear reasonably homogeneous, and therefore do not seem to reflect the bias introduced into the means. Bearing in mind the possibility of some distortion due to inadequate sampling, but finding no evidence of this in the variances to which the model is fitted, we conclude that the types of family represent reasonably adequate samples from the same population. Certainly, unless a low mean for the sample of MZ_A has restricted their variance to an appreciable extent, then the variance of this group provides no evidence for correlated environments.

The test for genotype-environment interaction yields the following correlations:

MZ_T (sexes pooled) $r_{41} = -0.3678$, $p = .02$

MZ_A (sexes pooled) $r_{38} = -0.0021$, *ns*

TABLE 13

AN ANALYSIS OF VARIANCE TO COMPARE THE FIVE MEANS IN TABLE 12

Source	*df*	*MS*	*F*
Between types	4	48.9489	2.43
Between families within types	99	20.1180	1.98*
Within families	94	10.1622	
Between types partitioned			
Sex difference	1	2.4870	
MZ_T versus MZ_A	1	190.6514	9.48*
Sex × MZ_T versus MZ_A	1	1.2950	
$MZ_{T\&A}$ versus DZ_T	1	1.3623	

*$p < .01$.

Thus there is evidence of a certain amount of GE_1 but not GE_2. Setting upper and lower bounds by the method previously described, GE falls within the range, 1.4590 to 0.1974. This negative correlation indicates that introvert genotypes are more susceptible to environmental influences than extravert genotypes, the latter being relatively impervious. This finding is, of course, fully consistent with Eysenck's (1960a) theory that the introvert is more conditionable than the extravert. There is a considerable amount of evidence for the theory, (Eysenck, 1960b), and our finding provides additional support. It would be of considerable interest to attempt to determine what kinds of environmental pressures are, in fact, contributing to this interaction by looking at the different effect on high and low genotypes of factors known to affect the trait measured by this test. Example 8, Table 4, illustrates the form such an investigation might take.

Fitting models. In spite of a certain amount of GE_1, we will fit the simple G and E model, and then by inspecting E_1 and the bounds for GE_1, see if the detected amount is of importance.

First we will fit the model G, E_1, E_2 to the mean values of the $\bar{V}_{FS}$ and $V_{\bar{F}S}$ listed for monozygotic twins in Table 11. The method is exactly as in the previous example, the matrix equations carrying the same coefficients.

$\hat{G} = 9.3192 \pm 1.8856$, $t_5\hat{G} = 4.94$, $p = .005$

$\hat{E}_1 = 7.3080 \pm 1.7669$, $t_5\hat{E}_1 = 4.16$, $p = .01$

$\hat{E}_2 = -2.4656 \pm 2.3558$, $t_5\hat{E}_2 = 1.05$, $p = .3$

The test for the model is $F_{1,4} = V(O\text{-}E)/V$ (error) $= 2.16$, which is clearly nonsignificant ($p = .2$). Thus the simple model provides an adequate description of the observations with both $\hat{G}_1$ and $\hat{E}_1$ highly significant. The negative $\hat{E}_2$ does not differ from zero. However, we will defer discussion of this component until later.

With $E_1 = 7.3080$ and $\widehat{GE}_1$ lying between 1.4590 and 0.1974 the amount of bias, although significant, is trivial. Allowing $\widehat{GE}_1$ to fall midway between its limits, it is only 0.8282, some 11% of E_1 and 6% of the total variation. Probably, we can safely ignore its effects in subsequent calculations. Fitting the full model G_1, G_2, E_1, and E_2, we obtain the following

estimates which have the significance shown.

$\hat{G}_1 = 16.2696 \pm 3.5286$,
$t_6\hat{G}_1 = 4.61$, $p = .004$

$\hat{G}_2 = -5.4667 \pm 4.4547$,
$t_6\hat{G}_2 = 1.23$, $p = .3$

$\hat{E}_1 = 8.3069 \pm 2.5970$,
$t_6\hat{E}_1 = 3.20$, $p = .02$

$\hat{E}_2 = -2.4656 \pm 3.5286$,
$t_6\hat{E}_2 = 0.70$, $p = .5$

Only $\hat{G}_1$ and $\hat{E}_1$ are significant, and $\hat{G}_2$ is negative, but not significantly. $\hat{E}_2$ is again non-significant ($p = .25$), but the emergence of both these two between-family components, $\hat{E}_2$ and $\hat{G}_2$, as negative requires an explanation. The equality of the total σ^2s gives $F_{2,4} = 6.18$, $p = .06$ indicating a failure of the model now that we have introduced the dizygotic twin sample into the estimation. This failure, it will be remembered, was also indicated by the test for heterogeneous variances given previously ($p = .1$), but does not suggest correlated environments because σ_T^2 tog/σ_T^2 ap was not significant and the simple G, E_1, E_2 model fit was clearly adequate. We suggest that the simple model has failed on the following counts:

1. The model does not fit adequately when dizygotic twins are included.
2. G_2 and E_2 are probably negative so that they cannot be equated to theoretical variances, which must always be positive in the linear model.
3. There is a certain amount of genotype-environment interaction.
4. G_1 is significantly different from zero; G_2 is not.

On genetical grounds, as we have seen from the discussion of gene action and the mating system, a large discrepancy between G_1 and G_2 is not possible. Failure due to Count 3 above cannot cause failure due to Count 1 or 2 and can only cause Count 4 if a large amount of GE_2 (see the section *Genotype-Environment Interaction*) is present. We have detected only GE_1. Counts 1, 2, and 4 therefore require an explanation in terms other than those previously suggested for failure of the model. The reason for difficulty, on introducing dizygotic twins is, of course, the very large σ_W^2 which implies a negative σ_B^2 for this group. The linear statistical model used in deriving the expectations of these components does not allow them to become negative unless the individuals within pairs are negatively correlated. This will occur with dizygotic twins if they react against each other in such a way as to develop opposite characteristics with respect to a trait. In doing this they will be reacting on a basis of differences due to G_1 as well as those due to E_1, whereas the same tendency in the monozygotic twins will only have E_1 effects to build upon. The negative covariance in the dizygotic twin pairs will therefore, be more pronounced than in the monozygotic twin pairs. This process could account for failure due to Counts 1, 2, and 4. This reaction of one twin to the other might have its origin in the intrauterine environment where one twin takes up a position favorable to development, and the other a less favorable one. Thus, if these positions are maintained, initial differences become accentuated. This phenomenon is termed "competition," and often takes place during the early part of the lives of many wild plants and animals. The runt, for example, in a litter of mammals is often the result of this kind of effect. A strong case may be argued for the intrauterine environment of twins producing strong competition (Burt, 1966; Burt & Howard, 1956) and differences in birth weight, which with competition may be quite pronounced, can result in the heavier twin assuming the more dominant role as Shields (1962) has shown. He found, and cited other studies which show that the leader twin is generally heavier at birth. He was not, however, able to show in his study that the heavier twin was more extravert, although in the MZ_T group alone there is a suggested association which fails to reach significance (Shields, 1962, see Table 20). The relevant information for the DZ_T group, which, on our hypothesis, would be expected to show the most pronounced effect is, unfortunately, not given. However, whether the birth weight is responsible for the initial differences in extraversion, or not, there is a strong association between leadership and extraversion (Shields, 1962, see Table 19). Moreover, the leadership pattern seems to develop continuously throughout the lives of the twins, their complementary roles becoming firmly established by adulthood. This leader-

TABLE 14

EXPECTATIONS OF VARIANCE COMPONENTS FOR THREE KINDS OF FAMILIES ACCORDING TO A GENETICAL MODEL SUITABLE FOR EXTRAVERSION

Type of family	Variance component	Genetical model			
		$G_1 = G_2 = \frac{1}{2}G$	E_1	E_2	$2CG_1$
MZ_T	$\bar{V}_F$	0	1	0	0
	$V_{\bar{F}}$	2	$\frac{1}{2}$	1	0
MZ_A	$\bar{V}_F$	0	1	1	0
	$V_{\bar{F}}$	2	$\frac{1}{2}$	$\frac{1}{2}$	0
DZ_T	$\bar{V}_F$	1	1	0	-1
	$V_{\bar{F}}$	$1\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$

ship effect which is not genotype-environment interaction or correlated environments would require an extension of the biometrical model. Before attempting this it would be necessary to detect the process by an unambiguous test. One such test might be afforded by a comparison of dizygotic twin groups reared together and apart. Also dizygotic twins might be compared with sibs who would be expected to show a similar but less pronounced effect due to the closer proximity of the twins both before and after birth. In Shields' data it is not possible to make these comparisons. A study by Portenier (1939) did however show the latter effect. Of a series of 12 personality measures, 9 showed smaller correlations for dizygotic twins than for sibs. This was particularly pronounced in the introversion score (dizygotic twins $r = -0.02$, sibs $r = 0.52$). The tests involved in Shields' and in Portenier's studies were, of course, not the same. Portenier's finding does, however, illustrate the sort of effect expected, and is included mainly to demonstrate an appropriate method.

With only the three types of twins it is not possible adequately to fit a suitable model. However, we can attempt some assessment of this negative correlation effect by making certain simplifying assumptions. It should be pointed out that the following assessment can only be tentative, and is included mainly to show the flexibility of the biometrical approach.

If we allow that H_R is small and there is little assortative mating, $G_1 = G_2 = \frac{1}{2}G$ is a good approximation. Then we can postulate a competition parameter CG_1 which describes that part of the covariance between dizygotic twins due to their genetic differences (G_1). That part due to E_1 cannot be allowed for and remains as a bias making E_1 too large and E_2 too small. We cannot allow for CE_1 because we cannot make any assumptions about the relative sizes of the between- and within-family environmental components, as we can for the genetic ones. The model is given in Table 14. The inverse of the matrix of coefficients of the normal equations is

$$\begin{bmatrix} 0.2375 & 0.0500 & -0.4000 & 0.2375 \\ 0.0500 & 0.8667 & -0.8000 & 0.8500 \\ -0.4000 & -0.8000 & 1.6000 & -1.2000 \\ 0.2375 & 0.8500 & -1.2000 & 1.8375 \end{bmatrix}$$

and the estimates are as follows:

$$\tfrac{1}{2}\hat{G} = 5.4014 \pm 1.3595, \qquad t_{6\frac{1}{2}\hat{G}} = 3.97,\ p = .01$$

$$\hat{E}_1 = 8.3069 \pm 2.5970, \qquad t_{6\hat{E}_1} = 3.20,\ p = .02$$

$$\hat{E}_2 = -2.4656 \pm 3.5286, \qquad t_{6\hat{E}_2} = 0.79,\ p = .50$$

$$2\widehat{CG}_1 = -10.8682 \pm 3.7815, \qquad t_{6\widehat{CG}_1} = 2.87,\ p = .04$$

The F test for the model is, of course, the same as before, and the same error variance is appropriate for obtaining the standard error of the estimates. The E_2 and E_1 still take the same errors. The competitive element is significant but our model only applies to the groups we have used and would seem to have little generality for the population at large. Heritabilities can be calculated, but would depend for their interpretation on whether or not the effect of CG was ignored.

Mill Hill Vocabulary Test

The next psychological measure chosen for analysis is the synonyms section of the Mill Hill Vocabulary Test. This test of verbal intelligence has two parts. In one, the subject chooses synonyms for the underlined word from six alternatives, and in the other is asked to define words. The two halves generally correlate quite highly (about 0.90 or better), and either half may be used separately to provide a shortened test. Unfortunately, the validity

of this test is not well established, and it is not known how it correlates with other measures of verbal intelligence. Certainly there are a number of cognitive levels implied in different vocabulary responses, and it seems unlikely that this test taps more than a small number of them. The author of the test (Raven, 1956) described it as a test of "acquired information" but this is misleading since verbal tests, this one included, show high heritabilities (Shields, 1962). Shields' (1962) gives data for 36 pairs of MZ_T and 40 pairs of MZ_A for this phenotype.

Analysis of variance and estimation of variance components. Significant between-family components were indicated for all the four groups and are given in Table 15.

Testing the assumptions. To see if the individuals in the four types of families have been drawn at random from the same population, we look for homogeneity of the means and variances listed in Table 16. An analysis of variance comparing the four means is given in Table 17. This analysis shows that there are differences between families and between types of families, the scores of monozygotic twins apart being significantly lower than for those reared together. There are, however, no sex differences in the mean so that we can again regard sexes as providing replication. This difference between twins reared together and apart clearly indicates that we are not sampling the same population. When we inspect the variances, heterogeneity, again, is clearly evident.

The $S_{max}^2/S_{min}^2 = 3.93$, $p = .01$ and an F ratio for $\sigma_T^2 ap/\sigma_T^2 tog = 2.12$, $p = .005$. Fortunately, it is possible to obtain a reasonably good estimate of the σ_T^2 for this test from the standardization data published in the instruction manual (Raven, 1956). A value of $\sigma_T^2 = 62.316$ was obtained and compared with the values in Table 16.

The F ratios given in Table 18 were obtained. All except the last F ratio are clearly significant, strongly suggesting inadequate sampling particularly for MZ_T. Inspection of the raw data indicates, in fact, that for the MZ_T males the lower 25% of the standardization sample is completely missing. The MZ_T are therefore a poor representation of available genotypes in the population. This restricted range is reflected fairly obviously in the differences in $V_{\bar{F}}$ for the together and apart groups in Table

TABLE 15

ESTIMATES OF VARIANCE COMPONENTS FOR TWO KINDS OF FAMILIES (MILL HILL VOCABULARY TEST)

Type of family	Variance component	Estimate	
		female	male
MZ_T	$V_{\bar{F}}$	17.2971	8.6119
	$\bar{V}_F$	4.6250	3.0419
MZ_A	$V_{\bar{F}}$	28.5834	34.3382
	$\bar{V}_F$	7.1400	10.6333

TABLE 16

TABLE OF MEANS AND VARIANCES FOR THE FOUR KINDS OF FAMILIES (MILL HILL VOCABULARY TEST)

Type of family	Mean	Variance
MZ_T		
female	18.6667	19.2908
male	19.5417	9.8243
MZ_A		
female	16.3000	31.6429
male	14.9667	38.6540

TABLE 17

AN ANALYSIS OF VARIANCE TO COMPARE THE FOUR MEANS IN TABLE 16

Source	*df*	*MS*	*F*
Between types of family	3	141.1953	2.85*
Between families within types	67	49.5312	7.24***
Within families	71	6.8380	
Between types of family partitioned			
Sex difference (S)	1	1.6695	
MZ_T versus MZ_A	1	383.0883	7.74**
S × MZ_T versus MZ_A	1	38.7687	

*$p < .05$.
**$p < .01$.
***$p < .001$.

TABLE 18

F RATIOS BETWEEN σ_T^2S FROM TABLE 16 AND σ_T^2 ESTIMATED FROM STANDARDIZATION DATA (MILL HILL VOCABULARY TEST)

Type of family	Sex	*F*
MZ_T	females	$F_{1000,48} = 3.23$**
	males	$F_{1000,24} = 6.34$**
MZ_A	females	$F_{1000,50} = 1.97$*
	males	$F_{1000,30} = 1.61$

*$p < .01$.
**$p < .001$.

TABLE 19

ESTIMATES OF VARIANCE COMPONENTS FOR TWO KINDS OF FAMILIES (DOMINOES INTELLIGENCE TEST)

Type of family	Variance component	Estimate	
		Female	Male
MZ_T	$V_{\bar{F}}$	58.1009	45.2546
	$\bar{V}_F$	24.8261	4.9091
MZ_A	$V_{\bar{F}}$	90.9275	54.7541
	$\bar{V}_F$	19.2500	19.3214

15. While we are not able to make use of these data, therefore, in the form of analysis used in this paper we can use them to demonstrate the importance of adequate testing of assumptions before proceeding with a complex analysis. If the values of the components in Table 15 are taken at face value, a very large $r_{bh,we}$, and, by implication $r_{wh,we}$, is indicated since $V_{\bar{F}}$ reflects σ_B^2 and

$$\sigma_B^2 \text{ ap} - \sigma_B^2 \text{ tog} = 2r_{bh,we}\sigma_{bh}\sigma_{we} - 2r_{bh,be}\sigma_{bh}\sigma_{be}$$

in the MAVA system.

We can, of course, still use these data to carry out a genotype-environment interaction test by correlating the means and differences of the MZ_T and MZ_A as described previously. The results are

$$MZ_T,\ r_{34} = -0.2758,\ p = .1$$
$$MZ_A,\ r_{38} = -0.2781,\ p = .07$$

indicating borderline significance for GE_1 and GE_2. The negative sign of the correlation indicates that environmental deviations are larger for the individuals with the lower IQs. These individuals seem, therefore, to be more at the mercy of the environment than those with higher IQs and, perhaps, reach their potential with less certainty. This fact would, of course, have important implications for educational practice suggesting that individuals at the lower end of the distribution need more careful nurturing than those at the higher end, if they are to develop fully their verbal intelligence potential.

However, the correlations observed here indicate that only 8% of the environmental variation arises from interactions with the genotype, and so the effect of differential education would be quite small. In view of the small effect and its dubious significance, it is impossible to regard these findings as any more than merely encouraging of further investigation.

We can use the one satisfactory group, male MZ_A, to estimate G and $E_1 + E_2 + GE_1 + GE_2$, and hence estimate the broad heritability for this test. Recalling the expectation for this group as $\sigma_B^2 = G$, and $\sigma_W^2 = E_1 + E_2 + GE_1 + GE_2$, and the broad heritability is

$$\hat{G}/(\widehat{E_1 + E_2 + GE_1 + GE_2}) = 0.7281$$

or about 73%.

Dominoes Intelligence Test

The final measure that Shields (1962) recorded is the Dominoes test score. In this test the subject is required to write the number of pips which should appear on a blank domino to complete the logical pattern formed by the other dominoes in the item. The test is like Raven's Progressive Matrices in that it involves a similar perceptual-cognitive task proceeding from easy to more difficult items. Shields' gives the reliability of the test as 0.92, and states that it correlates highly (0.86) with a general intelligence factor *g*. In view of this, and the fact that the test is timed to last only 20 minutes, it probably provides a quick, reliable measure of general intellectual and reasoning ability. For this phenotype 34 pairs of MZ_T and 38 pairs of MZ_A were available.

Analysis of variance and estimation of variance components. Estimating variance components led to Table 19.

Testing the assumptions. To test the adequacy of the sampling, an analysis of variance of the

TABLE 20

TABLE OF MEANS AND VARIANCES FOR THE TWO KINDS OF FAMILIES (DOMINOES INTELLIGENCE TEST)

Type of family	Mean	Variance
MZ_T		
female	25.8478	69.5097
male	32.3636	45.6710
MZ_A		
female	24.3333	98.8227
male	23.6786	62.7447

means in Table 20 was carried out and given in Table 21. We see there is evidence of inadequate sampling from the significance of the family-types item which, on further analysis is seen to result from the difference between groups reared together and apart. Again we may regard sexes as replicates. Among the variances $S_{max}^2/S_{min}^2 = 2.15$, $p = .1$, and $\sigma_T^2 ap/\sigma_T^2 tog = 1.39$, $p = 2 \times .05 = .1$, suggesting some heterogeneity. However, whereas for the Mill Till Test there was striking consistency of $\bar{V}_F$ and $V_{\bar{F}}$ between replicates (see Table 15), such is not the case for these data as the analysis of variance of the Vs in Table 19 shows (see Table 22).

Thus the heterogeneity of variance that initially appears as the result of correlated environments is more likely to be a sampling effect. Therefore, we conclude the sampling is inadequate and estimation from these data would be unreliable. However, if we drop the MZ_T males, who certainly contribute to the heterogeneity among the means and variances in Table 20, the analysis of variance of the means becomes as given in Table 23. While still retaining the significance of the between-family component, that between types has vanished, and $\sigma_T^2 ap/\sigma_T^2 tog = 1.22$, $p = .3$, no longer of borderline significance.

We conclude then that with these three groups there is adequate sampling and no evidence of correlated environments, and proceed with the analysis, remembering, of course, that our conclusions are limited to the female population.

Fitting the model. The tests for genotype-environment interaction yield for MZ_T, $r_{23} = -0.0878$ and for MZ_A, $r_{22} = -0.1051$, both nonsignificant values. With no correlated environments, no genotype-environment interaction and an adequate sample, we are justified in fitting a simple G, E_1, E_2 model.

The following values were obtained:

$$\hat{G} = 59.6070 \pm 18.8384, \quad t_3\hat{G} = 3.16, \; p = .05$$

$$\hat{E}_1 = 27.2200 \pm 17.6524, \quad t_3\hat{E}_1 = 1.54, \; p = .2$$

$$\hat{E}_2 = -10.3283 \pm 23.5365, \quad t_3\hat{E}_2 = 0.44, \; p = .06$$

Thus $\hat{E}_2$ is again negative and not significantly greater than zero, while only $\hat{G}$ is significant. Dropping E_2 from the model we obtain

$F_{2,2}$ model fit = 0.13, a nonsignificant deviation from the expected, and

$$\hat{G} = 54.4429 \pm 13.1420, \quad t_4\hat{G} = 4.14, \; p = .016$$

$$\hat{E}_1 = 22.0559 \pm 11.7546, \quad t_4\hat{E}_1 = 1.88, \; p = .15$$

These estimates are clearly subject to error in these data. However, we may use them to indicate broad heritability as 0.7117 or about 70%, which agrees well with estimates of heritability

TABLE 21

TABLE OF ANALYSIS OF VARIANCE OF THE FOUR MEANS IN TABLE 20

Source	*df*	*MS*	*F*
Between types of family	3	391.6670	2.94*
Between families within types	68	133.3500	7.07**
Within families	72	18.8611	
Between types of family partitioned			
Sex difference (S)	1	212.6414	1.59
MZ_T versus MZ_A	1	644.1356	4.83*
S × MZ_T versus MZ_A	1	317.4938	2.38

*p <.05.
**p <.001.

TABLE 22

ANALYSIS OF VARIANCE OF THE ESTIMATED VARIANCE COMPONENTS IN TABLE 19

Source	*df*	*MS*	*F*
MZ_T versus MZ_A; $V_{\bar{F}}$	1	447.7705	2.72
MZ_T versus MZ_A; $\bar{V}_F$	1	19.5187	
MZ_T versus MZ_A; $V_{\bar{F}} - \bar{V}_F$	1	5416.6108	20.83*
Replication	4	260.7312	

*p <.05.

TABLE 23

ANALYSIS OF VARIANCE OF MEANS IN TABLE 20 FOR MZ_T (FEMALES ONLY) AND MZ_A (BOTH SEXES)

Source	*df*	*MS*	*F*
Between types of family	2	47.9570	
Between families within types	57	143.1929	6.59*
Within families	60	21.7333	

*p <.001.

for IQ from other studies (Burt, 1966; Erlenmeyer-Kimling & Jarvick, 1963).

IQ (Burt's Group Test)

To illustrate the kind of analysis that can be undertaken using published intraclass correlations we will use those published by Burt (1966) for IQ and general school attainments (see Example 6, Table 4). For all but parental correlations, the IQ measurements were arrived at in the following way. Subjects were given a group test of intelligence containing both verbal and nonverbal items, and the results submitted to the teacher for comment. Where doubt was expressed the child was reexamined. The reliability of the group test was 0.97, and for a set of performance tests used occasionally in doubtful cases, 0.87. Both of these tests, which are described in detail elsewhere (Burt, 1921, 1933), were standardized by means of an individual test, the London Revision of the Terman-Binet, a broadly based test measuring many aspects of cognitive and intellectual ability. We have correlations based on 95 pairs of MZ_T, 53 pairs of MZ_A, 127 pairs of DZ_T, 264 pairs of FS_T, 151 pairs of FS_A, and 136 pairs of U_T.

To illustrate the extended analysis, we have taken estimates of the marital correlation (μ) and parent-offspring correlation ($r_{p.o}$) from Burt and Howard (1956). The numbers for these two correlations are not given. We have chosen Burt's correlations of the many possible sets presented by Erlenmeyer-Kimling and Jarvik (1963) because they are not only based on large numbers and many groups, but also result, largely, from the application of a single, highly reliable test. Moreover, they have been used by Burt and Howard (1956) to carry out calculation similar to ours. We believe our methods have certain advantages of scope and rigor over theirs which justify reanalysis, but it is useful to be able to check our broadly similar conclusions. The correlations we have chosen for reanalysis were taken from Table 4 in Burt's 1966 paper.

TABLE 24

CORRELATIONS AND SIMPLE GENETICAL MODEL FOR TWO KINDS OF FAMILIES (IQ)

Type of family	Correlation (r)	Model		
		G	E_1	E_2
MZ_T	0.92	1	0	1
MZ_A	0.87	1	0	0

TABLE 25

CORRELATIONS AND SIMPLE GENETICAL MODEL FOR THREE KINDS OF FAMILIES (IQ)

Type of family	Correlation (r)	Model			
		G_1	G_2	E_1	E_2
MZ_T	0.92	1	1	0	1
MZ_A	0.87	1	1	0	0
DZ_T	0.54	0	1	0	1

Analysis of correlations. Assuming no correlated environment, genotype-environment interaction and equality of E_1s and E_2s, we can fit the simple model given in Table 24 (remembering that r = fraction of total variance due to σ_B^2). The two equations in Table 24 may be solved for G and E_2, E_1 being obtained by subtraction of G and E_2 from the total 1.00. Thus,

$$\hat{G} = 0.87$$
$$\hat{E}_1 = 0.08$$
$$\hat{E}_2 = 0.05$$

suggesting a heritability of 0.87 or 87%.

Adding G_1 and G_2 to the model in the place of G, and using the correlations for DZ_T we can fit the model given in Table 25. Obtaining $\hat{E}_1$ by subtraction of $\hat{G}_1$, $\hat{G}_2$, and $\hat{E}_2$ from 1.00, we obtain the following estimates:

$$\hat{G}_1 = 0.38$$
$$\hat{G}_2 = 0.49$$
$$\hat{E}_1 = 0.08$$
$$\hat{E}_2 = 0.05$$

Using the alternative minimal set of data referred to in the previous discussion of this matter (see *Minimal Data*), and given together with the appropriate model in Table 26, we obtain G_1, G_2, E_1, and E_2. Again E_1 is obtained by subtraction:

$$\hat{G}_1 = 0.39$$
$$\hat{G}_2 = 0.43$$
$$\hat{E}_1 = 0.08$$
$$\hat{E}_2 = 0.10$$

The good agreement between the estimates obtained from the groups in Table 25 and those obtained from the groups in Table 26 strikingly confirms the adequacy of the latter groups as an alternative minimal set for a complete solution on the simple genetical model.

In both cases we notice $\hat{G}_2 > \hat{G}_1$ suggesting assortative mating is taking place. In fact, Burt and Howard (1956) found a marital correlation of $\mu = 0.3875$ for these data, indicating a considerable amount of assortative mating at the level of the phenotype, and used this value to investigate the mating system and gene action for this trait.

Before showing how our estimate of G_1 and G_2 may be used to carry out a similar investigation, it is worth considering how the best estimates of these parameters may be obtained from all the six groups available, rather than simply relying on an arbitrary selection of data as Burt and Howard have done. The method to be used is essentially that of least squares used in the reanalysis of Shields' (1962) data previously given; but in this case, because of the large numbers involved, we extend the analysis to take into account the different precisions of the correlations available. A weighted least-squares method will be used where each correlation is weighted according to the inverse of its sampling variance, a procedure leading to approximate maximum likelihood estimates of G_1, G_2, E_2, and E_1 and a chi-square test of the adequacy of the genetical model. The method only approximates to maximum likelihood since t is not normally distributed if large or based on a small sample, and we use t to estimate its own variance. However, the method certainly leads to a more comprehensive evaluation of the genetical model than hitherto attempted, and may be equally easily used with

TABLE 26

CORRELATIONS AND SIMPLE GENETICAL MODEL FOR THREE FAMILIES FORMING AN ALTERNATIVE SET TO THAT IN TABLE 25

Type of family	Correlation (r)	Model			
		G_1	G_2	E_1	E_2
MZ_T	0.92	1	1	0	1
FS_T	0.53	0	1	0	1
FS_A	0.44	0	1	0	0

analysis starting from raw σ_B^2s and σ_W^2s by using the inverse of the usual formulas for the variances of these parameters (Kempthorne, 1957) to provide the appropriate weights. For details of this approach to estimating genetical parameters, the reader is referred to Nelder (1960). The groups given in Table 27 may be used equating r's to the model in terms of G_1, G_2, and E_2, E_1 having been omitted from the model since it always carries a coefficient of zero and is obtained by subtraction as before.

The variance of r is obtained from $V(r) = (1 - r^2)^2/N$ so that the amount of information, $I(r)$ is approximately $1/V(r)$. Although, strictly this formula for $V(r)$ applies to product-moment correlations, and we are dealing with intraclass correlations, where family size is two, the more precise formula (Roberston, 1960) gives almost exactly the same values as the more familiar $(1 - r^2)^2/N$.

The normal equations are

$$\begin{bmatrix} 4914.6 & 4914.6 & 4016.1 \\ 4914.6 & 5811.1 & 4780.3 \\ 4016.1 & 4780.3 & 4938.5 \end{bmatrix} \begin{bmatrix} \hat{G}_1 \\ \hat{G}_2 \\ \hat{E}_2 \end{bmatrix} = \begin{bmatrix} 4476.507 \\ 4986.280 \\ 4145.087 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} \hat{G}_1 \\ G_2 \\ \hat{E}_2 \end{bmatrix} = \begin{bmatrix} 0.001209350 & -0.000967486 & -0.000046978 \\ -0.000967486 & 0.001552850 & -0.000716325 \\ -0.000046978 & -0.000716325 & 0.000934072 \end{bmatrix} \begin{bmatrix} 4476.507 \\ 4986.280 \\ 4145.087 \end{bmatrix}$$

where the 3×3 matrices are the information matrix and the variance-covariance matrix of the estimates, respectively.

The estimates are:

$$\hat{G}_1 = 0.39 \pm 0.03$$

$$\hat{G}_2 = 0.44 \pm 0.04$$

$$\hat{E}_1 = 0.08 \pm 0.01$$

$$\hat{E}_2 = 0.09 \pm 0.03$$

the standard errors being obtained as the square root of the appropriate leading diagonal term

TABLE 27

OBSERVED AND EXPECTED VALUES OF CORRELATIONS FOR SIX KINDS OF FAMILIES AND THE WEIGHTS $I(r)$ USED IN FITTING THE SIMPLE GENETICAL MODEL (IQ)

Type of family	Correlation (r)	Model			$V(r)$	$I(r)$	Expected (r)	(O-E)
		G_1	G_2	E_2				
MZ_T	0.92	1	1	1	0.000249	4016.1	0.9274	−0.0074
MZ_A	0.87	1	1	0	0.001113	898.5	0.8376	+0.0324
DZ_T	0.54	0	1	1	0.003945	253.5	0.5326	+0.0074
FS_T	0.53	0	1	1	0.001958	510.7	0.5326	−0.0026
FS_A	0.44	0	1	0	0.004305	232.2	0.4428	−0.0028
U_T	0.27	0	0	1	0.006321	158.2	0.0897	+0.1803

in the variance-covariance matrix, for example,

$$SE(\hat{G}_1) = \sqrt{(0.001209350)}$$
$$= 0.0347$$

The standard error of $\hat{E}_1$ is essentially that of r_{MZ_T}. If each $(O\text{-}E)^2$ is weighted by $I(r)$ then $\Sigma I(O\text{-}E)^2 = \chi_3^2$ which tests the adequacy of the genetical model

$$\Sigma I(O\text{-}E)^2 = 6.3251$$

Therefore,

$$\chi_3^2 = 6.3251,\ p = .10\text{–}.05,$$

indicating a borderline failure of the model. The reason for this failure is obviously the large value of $r_{U_T} = 0.27$, a direct estimate of E_2 estimated as 0.09 in the analysis using all the groups in Table 27. An approximate χ_1^2 for this single discrepancy is $0.1803^2 \times 158.2 = 5.1428$, $\chi_1^2 = 5.1428$, $p = .05\text{–}.01$, confirming the inadequacy of this correlation. Burt also rejected this correlation as atypical on the grounds that it probably reflected a placement effect due to the practice of fostering children with parents of similar estimated IQ. It is a feature of weighted estimation procedure that this correlation, in fact, introduces very little distortion into $\hat{G}_1$ and $\hat{G}_2$ because it contributes so little information compared with the other correlations, and yet its anomolous nature is detected by the chi-square test. A repeat of the estimation procedure omitting this correlation provides the following estimates:

$$\hat{G}_1 = 0.40 \pm 0.03 \quad c = 13.3 \quad p < .001$$
$$\hat{G}_2 = 0.47 \pm 0.04 \quad c = 11.8 \quad p < .001$$
$$\hat{E}_1 = 0.07 \pm 0.01 \quad c = 7.0 \quad p < .001$$
$$\hat{E}_2 = 0.06 \pm 0.03 \quad c = 2.0 \quad p = .025,$$

the significance of the parameter being shown by the normal deviate, c = parameter/SE and the model fit tested by

$$\chi_2^2 = \Sigma I(O\text{-}E)^2$$
$$= 0.2835,\quad p = .9\text{–}.8$$

which now shows an excellent fit for the model. The two χ^2s from this and the previous fitting may be compared approximately as

$$2\chi_3^2/3\chi_2^2 = F_{3,2}$$
$$= 12.09,\quad p = .10\text{–}.05$$

suggesting a significant increase in the adequacy of fit of the genetical model on omitting the correlation of U_T.

Analysis of gene action and mating system. We notice that $G_2 > G_1$, in all our estimations from these data, but that the difference does not reach significance $\hat{G}_2 - \hat{G}_1 = 0.07 \pm 0.07$, the error being obtained from the variance-covariance matrix with due allowance for the substantial correlation between $\hat{G}_1$ and $\hat{G}_2$. However, for these data we have independent evidence for assortative mating with the marital correlation $\mu = 0.3875$. Using this value of μ we will illustrate two methods for estimating H_R and D_R, the gene action parameters.

The first, an iterative method, uses A = Heritability $\times \mu$, in which our initial estimate of heritability is $(G_1 + G_2)/(G_1 + G_2 + E_1 + E_2) = 0.870$. Thus

$$A_1 = 0.87 \times 0.3875$$
$$= 0.3371$$

Substituting this value of A into the expectations for G_1 and G_2, we find that

$$G_1 = \tfrac{1}{4}D_R + \tfrac{3}{16}H_R$$
$$G_2 = \tfrac{1}{4}D_R + \tfrac{1}{16}H_R + \tfrac{1}{2}\{A/(1-A)\}D_R$$

which reduce to

$$G_1 = \tfrac{1}{4}D_R + \tfrac{3}{16}H_R = 0.40$$
$$G_2 = 0.504D_R + \tfrac{1}{16}H_R = 0.47$$

Solving these for D_R and H_R we find

$$\hat{D}_R = 0.808$$
$$\hat{H}_R = 1.184$$
$$\tfrac{1}{2}\{A/(1-A)\}D_R = 0.202.$$

These estimates provide a new estimate of heritability

$$= [\tfrac{1}{2}D_R + \tfrac{1}{2}\{A/(1-A)\}D_R]/$$
$$[\tfrac{1}{2}D_R + \tfrac{1}{4}H_R + \tfrac{1}{2}\{A/(1-A)\}D_R + E_1 + E_2] = 0.606$$

from which we estimate

$$A_2 = 0.606 \times 0.3875$$
$$= 0.2348$$

and the whole calculation is repeated with this new value of A.

Successive estimates of A were

$$A_1 = 0.337$$
$$A_2 = 0.235$$
$$A_3 = 0.264$$
$$A_4 = 0.260,$$

this last value being taken as the stable value, the difference between A_3 and A_4 being less than 2%.

The estimates obtained were

$$\hat{D}_R = 0.984$$
$$\hat{H}_R = 0.821$$
$$\tfrac{1}{2}\{A/(1-A)\}D_R = 0.173$$
$$H_R/D_R = 0.83$$

indicating a substantial amount of dominant gene action for this trait. These values are similar to those obtained by Burt and Howard (1956) who used $r_{p.o}$ and estimated $\frac{1}{2}D_R$, $\frac{1}{4}H_R$, and $\frac{1}{2}\{A/(1-A)\}D_R$ as percentages of total variance, concluding that only a small amount of dominance variation was present. While it is true that $\frac{1}{4}H_R$ is small (21%), it is the size of H_R compared with D_R which is relevant to a consideration of gene action, and this high ratio of 0.83 suggests the possibility of complete dominance.

On a single gene model and $u = v$

$$\sqrt{(H_R/D_R)} = h/d = 0.91$$

in this case. However, if $u \neq v$ this ratio may vary markedly from h/d, (Fisher, 1918; Mather, 1949).

The second method of estimation involves adding the statics $r_{p.o}/(1+\mu)$ to $\hat{G}_1$ and $\hat{G}_2$ to give three equations:

$$G_1 = \tfrac{1}{4}D_R + \tfrac{3}{16}H_R$$
$$G_2 = \tfrac{1}{4}D_R + \tfrac{1}{16}H_R + \tfrac{1}{2}\{A/(1-A)\}D_R$$
$$r_{p.o}/(1+\mu) = \tfrac{1}{4}D_R + \tfrac{1}{4}\{A/(1-A)\}D_R$$

The last being $\mathrm{Cov}_{p.o}/(1+\mu)$ when G_1 and G_2 are estimated from variance components. This gives

$$D_R = G_1 - 3G_2 + 6(r_{p.o})/(1+\mu)$$
$$H_R = 4G_1 + 4G_2 - 8(r_{p.o})/(1+\mu)$$
$$\tfrac{1}{2}\{A/(1-A)\}D_R = 1\tfrac{1}{2}G_2 - \tfrac{1}{2}G_1 - r_{p.o}/(1+\mu),$$

the errors of these estimates being given by

$$V(D_R) = V(G_1) + 9V(G_2) - 6\,\mathrm{Cov}(G_1G_2) + 36V[r_{p.o}/(1+\mu)]$$
$$V(H_R) = 16V(G_1) + 16V(G_2) + 32\,\mathrm{Cov}(G_1G_2) + 64V[r_{p.o}/(1+\mu)]$$
$$V[\tfrac{1}{2}\{A/(1-A)\}D_R] = 2\tfrac{1}{4}V(G_2) + \tfrac{1}{4}V(G_1) - 1\tfrac{1}{2}\,\mathrm{Cov}(G_1G_2) + V[r_{p.o}/(1+\mu)]$$

These expressions for the variance of the estimates are obtained by squaring the coefficients in the corresponding expressions for the estimates, and allowing cross-product terms to vanish in the case of $\hat{G}_1$ and $r_{p.o}/(1+\mu)$, $\hat{G}_2$, and $r_{p.o}/(1+\mu)$, where no correlation between the estimates exists, but leaving the coefficients for $\hat{G}_1$, $\hat{G}_2$ which are correlated.

The values of $V(\hat{G}_1)$, $V(\hat{G}_2)$ and $\mathrm{Cov}(\hat{G}_1\hat{G}_2)$ are obtained from the variance-covariance matrix provided by the estimation procedure. The $V[r_{p.o}/(1+\mu)]$ obtained by substituting into the formula for the variance of a ratio given in the example in the section *Neuroticism.*

For these data $r_{p.o} = 0.49$, so that

$$\hat{G}_1 = 0.396, \qquad V(\hat{G}_1) = 0.001209$$

$$\hat{G}_2 = 0.467, \qquad V(\hat{G}_2) = 0.001648$$

$$r_{p.o}/(1+\mu) = 0.353, \qquad V[r_{p.o}/(1+\mu)] = 0.000787$$

$$\mathrm{Cov}(\hat{G}_1\hat{G}_2) = -0.000961$$

Substituting in the three expressions for G_1, G_2, and $r_{p.o}/(1+\mu)$ we obtain

$$\hat{D}_R = 1.116 \pm 0.224, \qquad c = 4.98,\ p < .001$$

$$\hat{H}_R = 0.627 \pm 0.256, \qquad c = 2.45,\ p < .02$$

$$\tfrac{1}{2}\{A/(1-A)\}D_R = 0.149 \pm 0.074, \qquad c = 2.00,\ p < .05$$

thus showing a significant amount of dominance variation, although the value of H_R is slightly reduced in this procedure probably due to $r_{p.o}$ being biased by common environments. This conservative test for H_R indicates, therefore, that dominant gene action for IQ almost certainly exists. Unfortunately, the direction of dominance cannot be determined without the original data, although they contain the necessary groups of progeny from consanguineous matings and ordinary sib families (Burt & Howard, 1956). However, the significant negative skewness for the distribution of IQ scores which Burt (1963) has found certainly indicates that dominance is for high IQ, low being recessive. If this is indeed the case, the intuitively appealing idea that IQ has been subject to directional selection throughout man's evolutionary history would be strikingly confirmed.

Under assortative mating our best estimate of heritability is $2r_{p.o}/(1+\mu) = 0.706 \pm 0.009$, although this will probably be an overestimate due to common environments.

The broad heritability is given by

$$[\tfrac{1}{2}D_R + \tfrac{1}{4}H_R + \tfrac{1}{2}\{A/(1-A)\}D_R]/\sigma_T^2 = 0.863 \pm 0.003$$

The weakness of this estimation lies in the fact that we cannot test the adequacy of the model as is possible when we start with the raw variances. However, the σ_T^2s are probably homogeneous from evidence Burt gives elsewhere in his 1966 paper. In fact, an F test for MZ_A versus FS_T can be constructed and gives

$$F_{105,527} = 234.09/225.00 = 1.04, \qquad p = 2 \times .3 = .6,$$

strongly suggesting no correlated environments. The possibility of GE still remains. However, the probable absence of substantial genotype-environment interaction for Shields' two intelligence tests, and for the Newman, Freeman, and Holzinger (1937) data examined briefly in the final example (see Table 4, where a similar test to the Terman-Binet was used) make it at least plausible to suggest this source of bias is not likely to be very important either.

Educational Attainments

Educational attainments were measured by a group of scholastic tests (Burt, 1921, 1933) involving elementary reading, spelling, and arithmetic items. The items were similar to those in conventional school examinations.

Using correlations for the same groups and following exactly the same estimation procedure as in the previous example gave the following estimates:

$$\hat{G}_1 = 0.169 \pm 0.018, \quad c = 9.4, \quad p < .001$$

$$\hat{G}_2 = 0.509 \pm 0.050, \quad c = 10.0, \quad p < .001$$

$$\hat{E}_1 = 0.015 \pm 0.001, \quad c = 15.0, \quad p < .001$$

$$\hat{E}_2 = 0.307 \pm 0.052, \quad c = 5.9, \quad p < .001$$

and

$$\chi_2^2 = 1.1460, \quad p = .6$$

showing a satisfactory fit of the model to these data.

The value of c for $\hat{G}_2 - \hat{G}_1 = 6.12$, ($p < .001$) indicating that G_2 is significantly greater than G_1. However, a large discrepancy of this kind is simply not possible on a purely genetical model, and may well result from correlated environments ($r_{bh,be}$) or grossly inadequate sampling of the original population. Unfortunately, without recourse to the original raw data, it is not possible to decide which of these two possibilities is likely. If $r_{bh,be}$ is in fact responsible for producing this effect, it is interesting to note the marked contrast be-

tween this trait and IQ where neither E_2 nor $r_{bh,be}$ seem to be very important.

Perhaps, however, the main value of this example lies in demonstrating the inadequacy of the classical correlation approach when the assumptions implicit in simple genetical models are not met.

IQ (Several Tests)

To illustrate a single method we have chosen a small sample of 689 sib families of size three or more in order to calculate the mean within-family skewness from a vast population of family records published by Reed and Reed (1965). To do full justice to their unique data using the approaches we are suggesting would be a major undertaking, so we have restricted our analysis to a single method. Also we have limited ourselves to the range IQ 50–150 as being roughly comparable to the range of the other studies we have been considering, and to exclude the influence of single genes of very large effect. For this sample the mean skewness of the distribution of scores within families, $\bar{k}_3 = -257.326 \pm 121.023$, $c = 2.12$, $p = .03$, probably a significant result. The error is calculated from the variance of the 689 k_3 values obtained, weighted according to the degrees of freedom available for each k_3 (Fisher, Immer, & Tedin, 1932) and is only an approximate error. An alternative approach through the calculation of $g_1 = \bar{k}_3/\sqrt{(\sigma_W{}^3)}$ gives -0.2000 ± 0.0791, therefore, $c = 2.53$, $p = .01$ also suggesting the value of $\bar{k}_3$ is significantly less than zero. In the absence of genotype-environment interaction $\bar{k}_3 = -3\Sigma u^2v^2d^2h$. Our negative value indicates dominance for high IQ. Unfortunately we have no test for GE in these data, but two considerations cause us to place some confidence in our finding. Other examples (Dominoes and Mill Hill Vocabulary Test scores previously analyzed, and IQ scores in the final example) fail to reveal any significant GE for intelligence test scores, and our value $\bar{k}_3$ is consistent with the overall value of k_3 that characterizes the population from which our families were chosen.

Overall, $k_3 = -1077.102$, and $g_1 = -0.3386 \pm 0.0862$. Therefore $c = 3.93$, $p = .001$. If $u = v$, the k_3 should be equal to $4\bar{k}_3$ (Fisher, Immer, & Tedin, 1932). The agreement with this expectation in these data is striking. Moreover, since the effect of genotype-environment interaction is extremely unlikely to produce a value of $k_3 = 4\bar{k}_3$ even if present, we conclude that the negative value of $\bar{k}_3$ provides reasonable evidence for dominance for high IQ, and $k_3 = 4\bar{k}_3$ evidence for approximately equal frequencies for the genes controlling this trait. With equal gene frequencies $\sqrt{H_R/D_R}$ provides a valid estimate of the average level of dominance. Thus the value of this ratio (0.735) obtained from Burt's data may well indicate about the true level of dominance for IQ. It is worth pointing out that incomplete directional dominance is exactly what we would expect if high intelligence has been subject to a continual process of natural selection, as Reed (1965) has argued elsewhere. As in the example concerning the analysis of neuroticism, where a history of stabilizing selection was indicated, we are able, through the estimation of genetical parameters, to gain some insight into the past evolution of behavioral traits. This, to our mind, is the main justification for undertaking such an analysis.

In view of the very large sample in this study, the number of genes controlling IQ may reasonably be estimated using the range of IQ scores in the sample and the estimate of D_R from Burt's data. If n is the number of loci and there are equal allele frequencies ($u = v$) then $n = (\frac{1}{2}\,\text{Range})^2/D_R$. The range is 10–164 IQ points. While D_R, scaled up from $\sigma_T{}^2 = 1.00$ in Burt's data to 225 for Reed and Reeds' (1965) is 271, $n = 77^2/271 = 21.8$. Thus at least 22 loci would seem to be controlling IQ. This is probably a gross underestimate as n is lowered by the inequality of effects of the loci, and the range is probably underestimated due to ceiling effects in the tests. Clearly many loci are involved in determining individual differences in IQ.

IQ (Wechsler Intelligence Scale for Children)

Further evidence concerning dominance and the number of loci for IQ comes from a recent study by Schull and Neel (1965), reported by Spuhler (1967) on the effects of inbreeding on Japanese children. This showed strikingly that, for a number of tests, children resulting from consanguineous unions had lower scores than

those in a control group. This small (but highly significant) inbreeding depression provides very powerful evidence that dominant or interactive gene action is controlling these measures. We have singled out for consideration the IQ score derived from a Japanese version of the Wechsler Intelligence Scale for Children (WISC), an individual test of general intelligence covering a wide range of verbal and performance material. While the significant inbreeding depression is sufficient to establish that the genes controlling this trait show strong directional dominance, the size of the effect adds further information. Combining sexes and allowing for socioeconomic differences between the inbred and control groups (which Schull and Neel do by means of regression analysis) the estimated inbreeding depression is 3.70 IQ points (see Spuhler, 1967, Table 2). Remembering the formula in the section on gene action, $M_f = M_o - 2f\Sigma uvh$, where M_f is the mean of the inbred group, M_o that of the controls, and f the inbreeding coefficient, we can rearrange to give $\Sigma uvh = (M_o - M_f)/2f$. Then $(M_o - M_f)$ is the inbreeding depression.

In Schull and Neel's study the inbred group was the result of first-cousin marriages so that $f = \frac{1}{16}$. This gives

$$\Sigma uvh = 3.70 \times 8$$
$$= 29.6$$

If the estimate of H_R from the reanalysis of Burt's data may be taken as an estimate of H_R for this population, the number of genes showing directional dominance can be estimated from $n = 16(\Sigma uvh)^2/H_R$. Allowing that the use of H_R from another study is a doubtful procedure, it seems unlikely that its value would be entirely wrong. $\hat{H}_R = 0.627$ from Burt's data and scaled to σ_T^2 for the WISC IQ gives $\hat{H}_R = 141.075$.

$$n = 16 \times 29.6^2/141.075$$
$$= 99.33$$

Thus about 100 genes seem to be showing dominance for high IQ. This estimate is large and some caution should be adopted in placing too much reliance on it. Small biases in the study could result in considerable inflation of its value since the important item in the formula for n depends on $(M_o - M_f)^2$. However it is clear that directional dominance is demonstrated and that many genes are involved.

IQ (Stanford-Binet and Otis)

For our final example we have chosen to look at the IQ scores of the 19 pairs of MZ_A given in Newman, Freeman, and Holzinger's (1937) study to see if we can detect evidence of genotype-environment interaction. A more extensive analysis of their data is ruled out by the restricted range of the group we are considering, and by the absence of raw data on the DZ_T and MZ_T samples.

Two highly reliable and well-validated tests of general intelligence were used, the individual Stanford-Binet test, and the pioneer group test, the Otis Group Intelligence Scale.

Stanford-Binet test. Three tests of genotype-environment interaction were carried out.

1. For the 19 pairs of twins the correlation of $\frac{1}{2}$(sum) with $\frac{1}{2}$(difference) of their scores was calculated. This gave $r_{17} = -0.0996$, a non-significant result indicating absence of GE.

2. Cochran's test (Winer, 1962) for heterogeneity of variance was carried out. In this test $C = S^2\text{max}/\Sigma S^2$. In our case, each $\frac{1}{2}$(twin difference)^2 is an S^2 for 1 *df*.

It is these S^2s we subject to Cochran's test. Just as heterogeneity of variance within inbred lines denotes GE, so does heterogeneity of variance within twins (see *Tests for Genotype-Environment Interaction*). The maximum twin difference is 24 IQ points

$$\text{Therefore, } S^2\text{max} = \tfrac{1}{2}\,24^2$$
$$= 288$$
$$\text{Therefore, } C = 288/1061$$
$$= 0.2714$$

The .05 significance level for C is 0.3894, so we conclude there is, again, no evidence for GE.

3. Although with Tests 1 and 2 failing to reveal any effect, the test to be described is unlikely to produce a positive result, but it seems worth illustrating as part of a general approach to GE where it is present.

In the original study, each twin was rated for educational advantage relative to the other. Thus for each pair there is a difference score on this scale. This score seems from Newman,

Freeman, and Holzinger's (1937) analysis to largely account for the differences between twins. That is, $E_1 + E_2$ is almost entirely due to educational advantages. But it is worth asking if differences in educational advantage affect favorable genotypes to the same extent as unfavorable ones, that is, does genotype for IQ interact with educational advantage?

To examine this point, pairs of twins were divided into two groups on the basis of high or low genotype (i.e., according to their mean scores). They were then divided according to the size of difference in educational advantage (the nine largest difference scores being classed as "large educational difference"). A fourfold table of IQ differences was obtained (see Table 28). Analysis of variance was carried out on this table and gave the result in Table 29.

Thus the importance of educational differences in determining E_1 and E_2 is confirmed, but there is no suggestion that genotypes are differentially susceptible to environmental effects (genotype item), or that such differential sensitivity interacts with differences in educational advantage. This method could probably be extended to deal with a large range of causes of GE.

Otis IQ. Tests 1 and 2 were applied to the IQ scores obtained for the Otis Group Intelligence Scale.

1. $r_{17} = -0.1346$, a nonsignificant value.
2. Cochran's $C = 0.2088$, also a nonsignificant value. We conclude therefore, that there is no evidence of genotype-environment interaction in these data.

Discussion and Conclusions

The techniques of biometrical genetics described and illustrated in this paper can claim a number of advantages over rival methods of analysis.

One important advantage comes from preliminary testing of the basic assumptions underlying simple genetic and environmental models. These assumptions, discussed in the second main section, dealing with genotype-environment models, were tested for those examples where sufficient data was available, and the results, together with a summary of findings, are presented in Table 30. The assumptions, it will be recalled, are as follows: that all groups represent a random sample of genotypes and environments present in the population, that there is no correlation between genotype and environment and no influence of genotype-environment interaction.

TABLE 28

Table of Differences in IQ Between Each MZ_A Pair Classified According to a Favorable or Unfavorable Genotype and According to Difference in Education

Difference in educational advantage between each MZ_A pair	Genotype	
	Favorable	Unfavorable
Large	2, 17, 24, 7, 10	12, 15, 1, 19
Small	5, 8, 1, 6, 5	12, 4, 1, 1, 2, 9

Analysis of variance was carried out to test the adequacy of the sampling in each of the first four examples taken from Shields (1962), and allowed us to decide which groups were suitable for further analysis. Sampling was adequate for most of the groups in three of the measures but not, apparently, for some of those in the Mill Hill Vocabulary example. Further comparisons of group variances with the standardization variance confirmed our suspicion. In this case, inadequate sampling produced an effect similar to that of correlated environments, and further analysis would have been quite misleading. A test of heterogeneity of group variances was used to supplement the analysis of variance and heterogeneity of the mean variances, for groups reared together and apart, to indicate any correlation between genotype and environment. The latter test, applied in the first five examples, failed to reveal a significant correlation effect in two personality and three cognitive traits. Other features of the data in-

TABLE 29

Analysis of Variance of Environmental Deviations in Table 28

Source	*df*	*MS*	*F*
Genotypes (G)	1	0.04	
Educational difference (E)	1	41.7	4.97*
G×E	1	0.02	
Replication	15	8.39	

*$p < .05$.

TABLE 30

SUMMARY OF RESULTS AND FINDINGS FOR NINE EXAMPLES OF REANALYSIS OF THE DATA IN TABLE 4

Example		Main Findings				
Number	Phenotype	Correlated environments	Genotype-environment interaction	Heritability ± *SE*		Conclusions
				Broad	Narrow	
1	Neuroticism	none	none	54 ± 7%	54 ± 7%	Very simple genetical model adequate to explain data. Assortative mating indicated. Additive gene action only detected indicates intermediate expression of trait favored by natural selection. Common family environment unimportant.
2	Extraversion	none	Introvert genotype more modifiable than extravert genotype by within-family environment ($p < .02$)	67 ± 8%	—	Simple model not adequate as DZ_T appear to react against each other producing negative correlation. Some evidence effect is built upon genetic differences. Common family environment unimportant.
3	Mill Hill Vocabulary Test	none, but poor sampling mimics correlation effect	Suggestion that unfavorable genotypes are more influenced by environment than favorable ones (p.1–.05)	73 ± 12%	—	Poor sampling of genotypes vitiates attempt at more adequate analysis.
4	Dominoes Test	none	none	71 ± 7%	—	Very simple genetical model adequate. Common family environment unimportant.
6	Educational attainments	Positive correlation of heredity and environment between families indicated ($p < .001$)	—	Uncertain but probably less than 30%	—	Simple model probably not adequate, but sampling effects may be producing a false picture. No proper tests of assumptions possible without raw data. Assortative mating is indicated. Common family environment very important, and accentuated by effects of correlated environments.
5	IQ (Burt's Group Test)	none	—	86 ± 1%	71 ± 1%	Fairly simple model adequate to explain the inheritance of IQ. Assortative mating taking place.
7	IQ (Several tests)	—	—	—	—	Additive and dominant gene action detected with level of dominance 0.74.
8	IQ (WISC)	—	—	—	—	Dominance for high IQ for many genes (about 100) and gene frequencies, on average equal.
9	IQ (Stanford-Binet and Otis)	—	none	—	—	Gene action strongly suggests that IQ has been subject to considerable directional natural selection during man's evolutionary history. Unlike educational attainments, common family environment unimportant.

dicated correlation for educational attainments, as might be anticipated, but with no proper test of either sampling or correlation little reliance can be placed on this finding. The final assumption of no genotype-environment interaction was tested by correlating the sums and differences for MZ_T and MZ_A in the four examples taken from Shields (1962), and in the final example on IQ by a test of the heterogeneity of within-twin differences. Some interaction was indicated in the examples (see section *Extraversion* and the *Mill Hill Vocabulary Test*) but in neither case was the contribution to total variation important, being less than 10%. All these preliminary tests of assumptions allowed us to arrive at simple genetic and environmental models that were both realistic and parsimonious. In alternative approaches little or no attempt is made to test these assumptions with the result that we cannot be sure whether a particular model is appropriate or not. In the MAVA approach the omission of these tests may have serious consequences since the model is quite complex, and, if inappropriate, would

lead to a considerable distortion of the real situation.

Another important advantage of the biometrical approach comes from the method of estimation used in fitting the model finally chosen to represent the data. All other approaches use just sufficient groups to provide an exact fit, that is, the number of unknown parameters in the model equals the number of groups. Where more groups than parameters are available, as in the reanalysis of Burt's (1966) IQ data, this would result in an inefficient use of available information. When we consider the cost of data of this kind it seems imperative that it is used efficiently. The weighted least-squares or maximum likelihood procedure we describe uses all the available information in the data, and has the added advantage of leading to a chi-square test of "goodness of fit" of the model, providing a further test of the basic assumptions underlying the model. It also provides highly reliable standard errors for the estimated parameters without which further interpretation must remain uncertain. Using this method of estimation in the example reanalyzing Burt's data we showed that the U_T group was inconsistent with all other groups in the study, and thus justified its rejection. Subsequent estimation provided a strikingly good fit for a simple model yielding highly reliable estimates of G_1, G_2, E_1, and E_2. Finally, perhaps the most important contribution of the biometrical approach comes from the information it provides on the gene action controlling traits and the subsequent evolutionary implications that may be drawn from this. Adapting the approach of Fisher (1918) and Burt and Howard (1956) to benefit from the maximum likelihood estimation, we demonstrated two methods for estimating the gene action parameters D_R and H_R, allowing for the presence of assortative mating. For detecting the direction of dominance two methods were demonstrated, one using the shape of the distribution of scores within families, and the other the phenomenon of inbreeding depression. These methods, when applied to data on IQ (see sections *Burt's Group Test*, *Several Tests*, and *WISC*), all agreed in detecting a significant level of dominant gene action. The direction of dominance was for high IQ, indicating an evolutionary history of strong directional selection for this measure. The gene action parameters had the further value of allowing approximate values of the level of dominance compared with additive effects, the number of genes influencing the trait, the average gene frequencies, and the influence of assortative mating to be assessed.

Turning from the importance of the biometrical method to the findings, it is interesting to note that the inheritance of most of the psychological measures reanalyzed conform to a simple model. In view of pessimism, over the possible influence of correlated environments and genotype-environment interaction so often expressed in the psychological literature, it is reassuring to find they are by no means universal phenomena. The reasons why correlation effects are of little importance is not entirely clear to us, but may result from using tests having high test-retest reliability over long intervals. Such tests measure traits showing little dramatic change throughout long periods of the subject's life. However, these tests will, necessarily, measure aspects of subjects determined very early on, and may, therefore, reflect primarily genetic and prenatal and early postnatal influences. If this is so, many of the cultural factors, which would normally lead to correlated environments (Cattell, 1963), will produce little or no effect. This would also explain the frequent finding of the unimportance of common family environment (E_2). The absence of important genotype-environment interaction may also result from the use of tests with a high genetic component showing stability over long periods of time. A further factor may result, however, from the practice in test construction of aiming at a constant reliability throughout the range of the trait. Since error of measurement is included in twin differences, this would tend to minimize their heterogeneity. In view of the high heritabilities recorded in Table 30, and the known, or probable, reliabilities of the tests, much of E_1 may result from unreliability variance leaving little else environmental to interact with genotype. If we are correct, then measures not deriving from psychological tests might show a greater complexity of mode of inheritance as, for example, measures taken from experimental, physiological, or developmental psychology.

However, within the simple model the find-

ings for each trait showed important differences. For the section *Neuroticism*, assortative mating was indicated and only additive gene action, suggesting that the evolutionary history of this trait has involved natural selection for and intermediate optimum, either extreme of the trait being at a reproductive disadvantage. In contrast to this, examples concerned with IQ showed strong directional dominance for high expression, indicating that during man's evolution subjects with high IQ have been at a reproductive advantage, a situation to which man seems, recently, to be returning (Reed, 1965).

We note also that the high number of genes estimated to be controlling IQ (>22 and approximately 100) fully confirms that this trait is under polygenic control.

Extraversion, while not allowing a complete analysis, showed evidence of an interesting role taking effect. Shields (1962) discussed how one MZ_T twin assumes the extravert role of leader, and how this role develops throughout the lives of the twins. Our negative estimate of E_2 may result from this process not taking place when twins are separated, and the negative correlation item for DZ_T (CG_1) from genetic differences between DZ twins greatly intensifying this effect. Suggestions for the further investigation of this phenomenon are given in the example.

REFERENCES

BROADHURST, P. L., & JINKS, J. L. Stability and change in the inheritance of behavior in rats: A further analysis of statistics from a diallel cross. *Proceedings of the Royal Society, B*, 1966, **165**, 450–472.

BRUELL, J. H. Behavioral heterosis. In J. Hirsch (Ed.), *Behavior genetic analysis*. London: McGraw-Hill, 1967.

BURT, C. *Mental and scholastic tests*. London: P. S. King, 1921.

BURT, C. *Handbook of tests for use in schools*. London: Staples Press, 1933.

BURT, C. Is intelligence distributed normally? *British Journal of Statistical Psychology*, 1963, **16**, 175–190.

BURT, C. The genetic determination of differences in intelligence. *British Journal of Psycholgoy*, 1966, **57**, 137–153.

BURT, C., & HOWARD, M. The multifactorial theory of inheritance and its application to intelligence. *British Journal of Statistical Psychology*, 1956, **9**, 95–131.

CATTELL, R. B. The multiple abstract variance analysis equations and solutions for nature-nurture research on continuous variables. *Psychological Review*, 1960, **67**, 353–372.

CATTELL, R. B. The interaction of hereditary and environmental influences. *British Journal of Statistical Psychology*, 1963, **16**, 191–210.

CATTELL, R. B. Methodological and conceptual advances in evaluating hereditary and environmental influences and their interaction. In S. G. Vandenberg (Ed.), *Methods and goals in human behavior genetics*. New York: Academic Press, 1965.

DICKINSON, A. G., & JINKS, J. L. A generalized analysis of diallel crosses. *Genetics*, 1956, **41**, 65–78.

ERLENMEYER-KIMLING, L., & JARVIK, L. F. Genetics and intelligence: A review. *Science*, 1963, **149**, 320–321.

EYSENCK, H. J. Levels of personality, constitutional factors, and social influences: An experimental approach. *International Journal of Social Psychiatry*, 1960, **6**, 12–24. (a)

EYSENCK, H. J. *The structure of human personality*. London: Methuen, 1960. (b)

FISHER, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society (Edinburgh)*, 1918, **52**, 399–433.

FISHER, R. A., IMMER, F. R., & TEDIN, O. The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. *Genetics*, 1932, **17**, 107–124.

FULKER, D. W. Mating speed in male *Drosphila melanogaster*: A psychogenetic analysis. *Science*, 1966 **153**, 203–205.

GOTTESMAN, I. I. Personality and natural selection. In S. G. Vandenberg (Ed.), *Methods and goals in human behavior genetics*. New York: Academic Press, 1965.

HOLZINGER, K. J. The relative effect of nature and nurture influences on twin differences. *Journal of Educational Psychology*, 1929, **20**, 245–248.

HUNTLEY, R. M. C. Heritability of intelligence. In J. E. Meade & A. S. Parkes (Eds.), *Genetics and environmental factors in human ability*. London: Oliver & Boyd Ltd., 1966.

HUSÉN, T. *Psychological twin research*. Vol. 1. *A methodological study*. Stockholm: Almqvist and Wiksell, 1959.

JENSEN, A. R. Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins. *Proceedings of the National Academy of Sciences*, 1967, **58**, 149–156.

KEMPTHORNE, O. *An introduction to genetic statistics*. New York: Wiley, 1957.

LOEHLIN, J. C. Some methodological problems in Cattell's Multiple Abstract Variance Analysis. *Psychological Review*, 1965, **72**, 156–161.

MATHER, K. *Biometrical genetics: The study of continuous variation*. London: Methuen and Co., 1949.

MATHER, K. The genetical structure of populations. *Symposium of the Society for Experimental Biology*, 1953, **7**, 66–95.

MATHER, K., & MORLEY-JONES, R. Interaction of genotype and environment in continuous variation. I. Descriptions. *Biometrics*, 1958, **14**, 343–359.

NEEL, J. V., & SCHULL, W. J. *Human heredity*. Chicago: University of Chicago Press, 1954.

NELDER, J. A. The estimation of variance components in certain types of experiment on quantitative genetics. In O. Kempthorne (Ed.), *Biometrical genetics*. London: Pergamon, 1960.

NEWMAN, H. H., FREEMAN, F. N., & HOLZINGER, K. J. *Twins: A study of heredity and environment*. Chicago: University of Chicago Press, 1937.

NICHOLS, R. C. The National Merit twin study. In S. G. Vandenberg (Ed.), *Methods and goals in human behavior genetics*. New York: Academic Press, 1965.

PORTENIER, L. Twinning as a factor influencing personality. *Journal of Educational Psychology*, 1939, **30**, 542–547.

RAVEN, J. C. *Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales*. London: H. K. Lewis, 1956. (Reprinted 1962).

REED, S. C. The evolution of human intelligence. *American Scientist*, 1965, **53**, 317–326.

REED, E. W., & Reed, S. C. *Mental retardation: A family study*. London: W. B. Saunders Co., 1965.

ROBERTS, R. C. Implications of behavior genetics for genetics. In J. Hirsch (Ed.), *Behavior-genetic analysis*. London: McGraw-Hill, 1967.

ROBERTSON, A. Experimental design on the measurement of heritabilities and genetic correlations. In O. Kempthorne (Ed.), *Biometrical genetics*. London: Pergamon, 1960.

SHIELDS, J. *Monozygotic twins*. Oxford: Oxford University Press, 1962.

SCHULL, W. J., & NEEL, J. V. *The effects of inbreeding on Japanese children*. New York: Harper & Row, 1965.

SPUHLER, J. N. Behavior and mating patterns in human populations. In J. N. Spuhler (Ed.), *Genetic diversity and human behavior*. Chicago: Aldine Publishing Company, 1967.

VANDENBERG, S. G. Contributions of twin research to psychology. *Psychological Bulletin*, 1966, **66**, 327–352.

WINER, B. J. *Statistical principles in experimental design*. London: McGraw-Hill, 1962.

WRIGHT, S. Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics*, 1921, **6**, 144–161.

WRIGHT, S. The genetic structure of populations. *Annals of Eugenics*, 1951, **15**, 323–354.

(Received May 16, 1969)