Am J Hum Genet 26:331–359, 1974

Analysis of Family Resemblance. II. A Linear Model for Familial Correlation

D. C. Rao,[1] N. E. Morton, and S. Yee

Wright [1] developed path analysis to describe familial correlation under a linear model. This remains its principal use, despite extension to nonfamilial data and interactive systems [2, 3]. His seminal paper on the intelligence quotient has been surprisingly neglected by human geneticists, perhaps because values had to be assumed for some unknown parameters [4]. Recently this indeterminacy was resolved by combining path analysis with the concept of an index and a theory of hypothesis testing [5]. In this form it appears to be a powerful tool to analyze family resemblance and group differences in man. Since our work depends on path analysis, which, like so much in population genetics, is the creation of Sewall Wright, we take pleasure in dedicating this paper to him.

THEORY

In this section we describe the necessary theories of estimation and tests of hypotheses (e.g., see [6]) which can be used for any model; these will be used for the models we develop in later sections.

The traditional approach to familial correlation is to estimate heritability under assumptions that are not tested. A more fruitful alternative emphasizes tests of hypotheses. For a phenotype influenced by familial environment and/or subject to assortative mating, in an organism where opportunities for critical experimentation are severely limited, it would be reckless to suppose that any particular model is true, whereas it may be feasible to show that a particular model is false.

Our interest in familial correlation is to test hypotheses for which we require a distribution theory. Let r_N be an estimated intra- or interclass correlation based on N pairs of observations, and let ρ be its expected value. Fisher [7] showed that for an interclass correlation the bias-corrected z transform,

$$z = \frac{1}{2} \ln \left(\frac{1 + r_N}{1 - r_N} \right) - \frac{\rho}{2(N-1)},$$

is distributed almost normally with mean $\bar{z} = \frac{1}{2} \ln [(1+\rho)/(1-\rho)]$ and vari-

Received July 2, 1973; revised November 1, 1973.

PGL paper no. 112. This work was supported by grants GM 17173 and HD 06003 from the U.S. National Institutes of Health.

[1] All authors: Population Genetics Laboratory, University of Hawaii, Honolulu, Hawaii 96822.

ance $\sigma_z^2 = 1/(N - 3)$. For an intraclass correlation, the bias-corrected z transform is

$$z = \frac{1}{2} \ln \left(\frac{1 + r_N}{1 - r_N} \right) + \frac{1}{2} \ln \frac{N}{N - 1}$$

with mean $\bar{z}$ and variance $\sigma_z^2 = 1/(N - 3/2)$. If $z_1, z_2, \ldots, z_m$ are all independent, each having an approximately normal distribution, the logarithm of the joint likelihood may be written (ignoring terms that are independent of parameters) as $\ln L \propto -\chi^2/2$, where

$$\chi^2 = \sum_{i=1}^{m} (z_i - \bar{z}_i)^2 / \sigma_{z_i}^2 \tag{1}$$

is a χ^2 based on m correlations, with degrees of freedom equal to $m - \kappa$, where κ is the number of parameters estimated from the data. The χ^2 is a function of the ρ_i, which in turn are functions of a set of parameters. If λ_r, λ_s are two such parameters, the maximum-likelihood scores, neglecting $(\partial\rho/\partial\lambda_r)/[2(N_i - 1)]$ for interclass, are given by

$$u_{\lambda_r} = \frac{\partial \ln L}{\partial \lambda_r} = \sum_{i=1}^{m} (z_i - \bar{z}_i) \left(\frac{\partial \bar{z}_i}{\partial \lambda_r} \right) / \sigma_{z_i}^2$$

and

$$K_{\lambda_r \lambda_s} = -E \left(\frac{\partial^2 \ln L}{\partial \lambda_r \partial \lambda_s} \right) = \sum_{i=1}^{m} \left(\frac{\partial \bar{z}_i}{\partial \lambda_r} \right) \left(\frac{\partial \bar{z}_i}{\partial \lambda_s} \right) / \sigma_{z_i}^2.$$

The likelihood may be maximized, or the χ^2 minimized, by Newton-Raphson iteration. If λ is a vector of trial values for κ of the m parameters and λ_n is the vector at the nth iteration, then $\lambda_{n+1} = \lambda_n + UK^{-1}|\lambda_n$ leads to the maximum-likelihood values of λ, subject to a null hypothesis about the $m - \kappa$ other parameters.

In this way, one can estimate κ ($\leqslant m$) parameters. If $\chi^2_{m-\kappa}$ and $\chi^2_{m-\kappa-\omega}$ are two values of χ^2 (eq. 1), based on estimation of κ and $\kappa + \omega$ parameters from the data, where each of the κ parameters is included in the $\kappa + \omega$ parameters,

$$\chi^2_\omega = \chi^2_{m-\kappa} - \chi^2_{m-\kappa-\omega} \tag{2}$$

is a χ^2 on a null hypothesis about the ω parameters.

THE GENERAL MODEL

Suppose that J is an estimate of a cause K for a phenotype Y. Then J is called an *index* of K [5]. If there are n different estimates, say $J_1, \ldots, J_n$, they may be combined by the discriminant

$$J = \sum_{i=1}^{n} B_i J_i,$$

where B_i maximizes the multiple correlation of J_i with Y. We shall be concerned with the index (I) of family environment and the index (M) of race. In both cases the relations between estimates, causes, and effects are assumed linear, with the index determined by a single cause and a random error, $J = K + \epsilon$. The concept of an index makes it unnecessary (as it is undesirable) to assume that common environment or race is known precisely. Environment common to siblings reared together, usually including characteristics of the neighborhood, school, and social class, is called *home, family,* or *common,* interchangeably.

The phenotype Y of a child is determined linearly by midparent genotype (G), segregation from the midparent (S), family environment (C), race (N), and random environment (E) (see [5, fig. 1]). Thus, race is separated from the residual genotype, and errors of phenotype estimation are pooled with random environment. The phenotype of an adult is determined by the same factors as child's phenotype, but the effects of genotype and common environment may be different where common environment refers to the conditions under which the adult is rearing his family, not the home environment he experienced as a child (see table 1).

TABLE 1

VARIABLES IN THE GENERAL MODEL

Variable	Directly Observed	Description
Within race:		
X	Yes	The phenotype of the first of a pair of individuals; parent's phenotype in offspring-parent pairs; natural sib's phenotype in natural-adopted sib pairs
Y	Yes	The phenotype of the other member of the pair
C	No	Common environment; for parents, the home environment provided to children they rear
I	Yes	The index of common environment (estimate of C)
G	No	Midparent genotype
E	No	Random environment
Among races:		
N	No	The race of an individual
M	Yes	The racial index of an individual (estimate of N)
W	No	Midparent race

NOTE.—The subscript T denotes a factor common to monozygous twins (G_T = their genotype, N_T = their race, etc.); the subscript P denotes a factor acting on a parent but not directly on his child (G_P = parent's genotype, etc.).

Most of the paths are between causes and effects, represented by single-headed arrows and taking values between 0 and 1 (table 2). However, four pairs of factors are assumed correlated in a more ambiguous way, and these paths, which are between -1 and $+1$, are indicated by double-headed arrows. There may be a marital correlation m between the parental genotypes and w between the parental races; midparent genotype may be correlated to an extent r with common environ-

TABLE 2

PATH COEFFICIENTS OF THE GENERAL MODEL

Source and Symbol	Definition
Within races:	
c	Effect of common environment on child's phenotype
ck	Effect of common environment on parent's (adult) phenotype
h	Effect of genotype on child's phenotype (square root of "heritability")
hz	Effect of genotype on parent's (adult) phenotype
i	Effect of common environment on index
m	Correlation between parental genotypes
r	Correlation between midparent genotype and common environment
b	Correlation between common environments of half-sibs reared apart
Among races:	
t	Effect of race on racial index
q	Effect of race on phenotype
p	Correlation between midparent race and common environment
w	Correlation between parental races

NOTE.—The path from midparent genotype to child's phenotype is $g = h\sqrt{(1+m)/2}$. The path from midparent race to index of midparent race $= t\sqrt{(1+w)/(1+t^2w)}$. The path from midparent race to race of individual is $n = \sqrt{(1+w)/2}$.

ment, and there may be a correlation p between midparent race and common environment. For adopted children, r and p are assumed to be nullified. In the special case of half-sibs reared apart by their own parents, the home environments may be correlated to an extent b. Otherwise, the home environments of individuals reared apart are assumed to be uncorrelated. As a special case of the assumption that relations among estimates, causes, and effects are linear, all interactions are considered negligible, including dominance, epistasis, and gene-environment interactions. The effect of common environment is assumed to be the same for singletons and twins, both dizygous and monozygous, for half-sibs and full sibs reared together, and even for adopted and true children. Failure of these assumptions can in principle be detected but not always differentiated. Thus, epistasis and environment common to twins would both inflate the monozygous correlation, but could be distinguished by comparison of monozygous twins reared together and apart. However, common environment and gene-environment interaction could not be differentiated.

With more than two ancestral races, they may be ordered by their phenotypic means. Then the linearity assumption scores the race of hybrid individuals from their admixture proportions, $M = \sum_i X_i P_i$, where M is the racial index, X_i is the phenotypic mean of the ith ancestral race, and P_i is the admixture proportion. In the more common case of only two ancestral races, the racial index reduces to the admixture proportion.

Variance Components

Heritability (h^2) is defined as the proportion of the phenotypic variance due to additive genetic differences. It may be partitioned into a part due to variation

among midparental genotypes (g^2) and a remainder (s^2) due to segregation from the midparent, or $h^2 = g^2 + s^2$. By using the midparent in this way, we are implicitly accepting the polygenic model of an indefinitely large number of additive genetic factors under which the progeny distribution is determined by the midparent breeding value independently of the genotypic variance between parents. Positive assortative mating inflates the midparent effect at the expense of segregating factors, since $g^2 = h^2(1 + m)/2$ and $s^2 = h^2(1 - m)/2$. Similarly, race is taken as polygenic, which corresponds more closely to the Latin American convention that full sibs may be of different phenotypic race than to the North American superstition that race is completely determined by one or both parents. The proportion of the phenotypic variance due to race is q^2, which may be partitioned into a part due to midparent race (q^2n^2) and a residual, $q^2(1 - n^2)$, due to segregation from the midparent race, with $q^2n^2 = q^2[(1 + w)/2]$ and $q^2(1 - n^2) = q^2[(1 - w)/2]$. The null hypothesis that racial phenotypic differences are entirely environmental corresponds to $q = 0$, which can be tested if there is racial variation and the racial index is estimated.

Phenotypic standard deviations of children reared by their own parents and adopted children are in the ratio

$$\theta = 1/\sqrt{1 - 2(grc + qnpc)}.$$

It would be hard in practice to distinguish this from truncated placement, whereby children with extreme phenotypes are less likely to be put out for adoption.

Sometimes an estimate is made of the "attenuation" due to errors in measuring the phenotype. If α is this proportion of the variance, a correction for attenuation can be made by dividing all the systematic components of variation by $1 - \alpha$ and assigning to random environment the complement of the resulting sum (table 3).

TABLE 3

VARIANCE COMPONENTS AS FRACTIONS OF TOTAL PHENOTYPIC VARIANCE OF CHILDREN REARED BY THEIR OWN PARENTS (ATTENUATION CORRECTION FACTOR = PROPORTION OF PHENOTYPIC VARIANCE ATTRIBUTED TO ERRORS OF MEASUREMENT = α)

Source	Crude Variances	Variances Corrected for Attenuation
Within races:		
Random environment	e^2	$(e^2 - \alpha)/(1 - \alpha)$
Heritability	h^2	$h^2/(1 - \alpha)$
Common environment	c^2	$c^2/(1 - \alpha)$
Gene-environment covariance	$2grc$	$2grc/(1 - \alpha)$
Among races:		
Race	q^2	$q^2/(1 - \alpha)$
Race-environment covariance	$2qnpc$	$2qnpc/(1 - \alpha)$
Total	1	1

Standard errors of the variance component estimates are obtained as follows: let

$$\hat{V} = V(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_\kappa, \theta_{\kappa+1}, \ldots, \theta_{\kappa+\omega})$$

be an estimated variance component as a function of $\kappa + \omega$ parameters, of which only κ are estimated. Then the standard error of $\hat{V}$ is estimated by

$$\hat{\sigma}_{\hat{V}} = \sqrt{\underset{\sim}{\delta}' K^{-1} \underset{\sim}{\delta}},$$

where

$$\underset{\sim}{\delta}' = (\delta_1, \ldots, \delta_\kappa) : \delta_i = \left(\frac{\partial \hat{V}}{\partial \theta_i} \right) | \theta_i = \hat{\theta}_i, i = 1, 2, \ldots, \kappa$$

and K is the information matrix of $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_\kappa$. Tests of significance could in principle be made from these standard errors. However, estimates of variance components approach normality even more slowly than estimates of path coefficients, and so significance tests are better made from the likelihood equation.

SPECIFIC MODELS

Racial paths vanish for a study within racial groups. In presentation of the specific models below, the racial terms are enclosed in braces. We consider monozygous twins, full sibs, half-sibs, parent-offspring, and unrelated pairs, the members of which are reared together or apart. In the latter event, both may be reared by random parents, or one may be reared by his own parents and the other adopted randomly. Some investigators fail to determine zygosity of like-sexed twins, and so this case is treated separately. The various possibilities can be identified by a three-letter code, in which the first two letters specify the relation and the third letter indicates communality of the environment (figs. 1–16). We shall denote the phenotypes of the two individuals involved by X and Y. If X and Y are phenotypes in different generations, X is assigned to the older generation. Factors back of Y not common to X are designated by primes: N', M', G', etc. (figs. 1–16). Pairs of sibs or half-sibs reared together by unrelated foster parents have not been considered because (except perhaps for dizygous twins) they are likely to be of uncertain paternity or to show effects of the home provided by their true parents before adoption.

A correlation coefficient will be denoted by a five-letter code, in which the first three letters specify the relation and communality of the environment and the last two letters specify the variables (MZTXY, SSPXI, etc.).

Monozygous Twins Reared Together by Their Parents (MZT) (Fig. 1)

The most critical assumptions are: (1) phenotypic similarity due to common prenatal and postnatal environment is no greater or less than for ordinary siblings; and (2) dominance, epistasis, and gene-environment interactions are negligible. Failure of either assumption tends to give spuriously high estimates of heritability, an error that may in principle be detected by a goodness-of-fit test against other

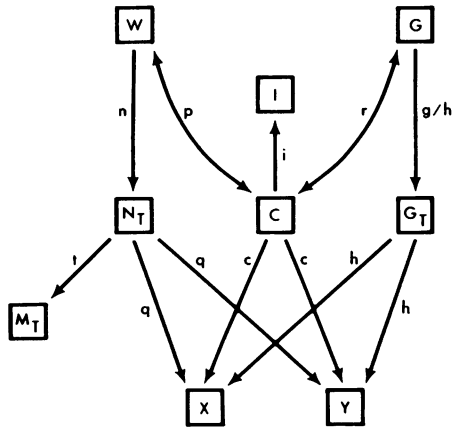

FIG. 1.—MZT. Monozygous twins reared together by their parents. G_T, N_T, and M_T denote the genotype, race, and racial index of the twin pair.

pairs of relatives. From figure 1 the expected correlation of MZ twins reared together is

$$\rho_{XY} = c^2 + h^2 + 2grc + \{q^2 + 2\,qnpc\}.$$

Monozygous Twins Reared Apart, One by True Parents, the Other by Random Foster Parents (MZP) (Fig. 2)

It is assumed that none of the phenotypic similarity is due to prenatal or postnatal environment shared before adoption, and that the adopted twin is placed in a random foster home. From figure 2 their correlation is

$$\rho_{XY} = (h^2 + grc + \{q^2 + qnpc\})\,\theta.$$

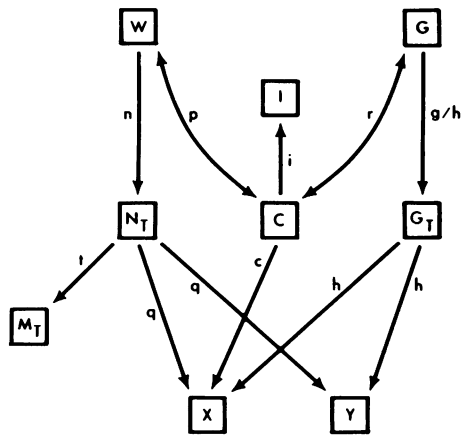

FIG. 2.—MZP. Monozygous twins reared apart, X by true parents and Y by random foster parents. Not shown are C' and I', the common environment and index of the foster home which reared Y.

Unless arteriovenous anastomosis is a relevant factor, failure of either assumption probably tends to inflate the correlation, giving a spuriously high estimate of heritability. An important test of random placement is the correlation between indices, $\rho_{II'} = 0$.

The principal limitation of this relationship is its rarity. It would be useful in a society which regularly placed for adoption one member of each set of twins.

Monozygous Twins Reared Apart by Random Foster Parents (MZA) (Fig. 3)

The critical assumptions and their test are the same as for MZP. The correlation between twins is

$$\rho_{XY} = (h^2 + \{q^2\})\,\theta^2.$$

Suitable data are extremely rare.

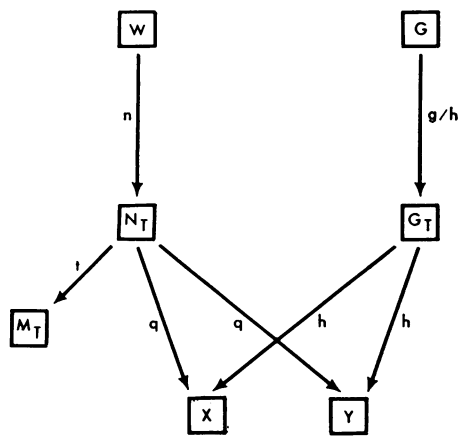

FIG. 3.—MZA. Monozygous twins reared apart by random foster parents. Common environments and indices of the two foster homes, C and I for X and C' and I' for Y, are not shown.

Monozygous Twins Reared Together by Random Foster Parents (MZF) (Fig. 4)

This is another rare type. Environment shared before adoption is neglected. Random placement is tested by the correlation between indices of true and foster parents, $\rho_{II'} = 0$. The twin correlation is

$$\rho_{XY} = (c^2 + h^2 + \{q^2\})\,\theta^2.$$

Full Sibs Reared Together by Their Parents (SST) (Fig. 5)

Under the model, dizygous twins are no more alike than ordinary siblings, which should be tested. The correlation is

$$\rho_{XY} = c^2 + g^2 + 2grc + \{q^2n^2 + 2qnpc\}.$$

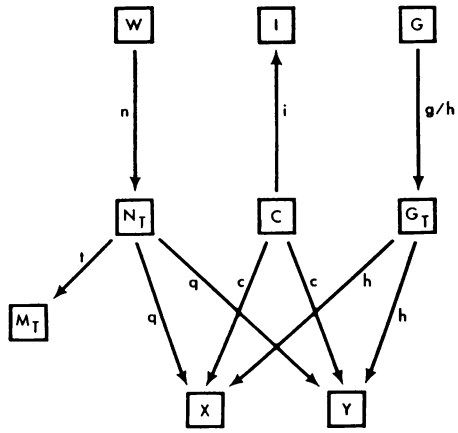

FIG. 4.—MZF. Monozygous twins reared together by random foster parents. *C* and *I* refer to the foster home.

Full Sibs Reared Apart, One by True Parents, the Other by Random Foster Parents (SSP) (Fig. 6)

This is analogous to MZP, with a genetic covariance only half as great. It is assumed that none of the phenotypic similarity is due to shared prenatal and postnatal environment before adoption, and that the adopted sib is placed in a random foster home. Such pairs may be frequent in societies which encourage adoption (as in the *hanai* system of Polynesians and Micronesians), but the assumptions would be violated if the child were adopted by close relatives or

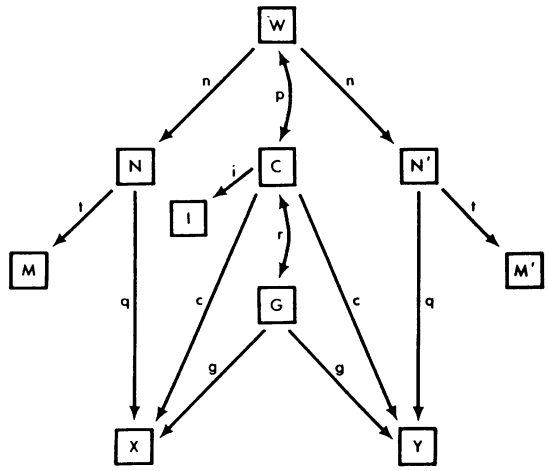

FIG. 5.—SST. Full sibs reared together by their parents. Conventionally, *N'* and *M'* denote the race and its index for *Y*, like *N* and *M* for *X*.

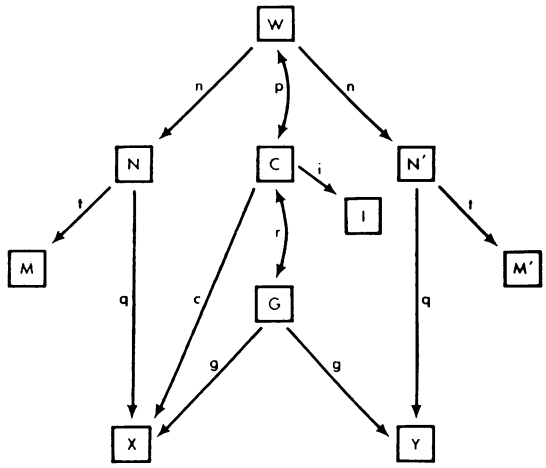

FIG. 6.—SSP. Full sibs reared apart, X by true parents and Y by random foster parents. For Y, the common environment C' and index I' of the foster home are not shown.

continued to be influenced by his true parents. The expected phenotypic correlation is

$$\rho_{XY} = (g^2 + grc + \{q^2n^2 + qnpc\})\ \theta.$$

Full Sibs Reared Apart by Random Foster Parents (SSA) (Fig. 7)

This is analogous to MZA, and the critical assumptions and their test are the same as above. The correlation is

$$\rho_{XY} = (g^2 + \{q^2n^2\})\ \theta^2.$$

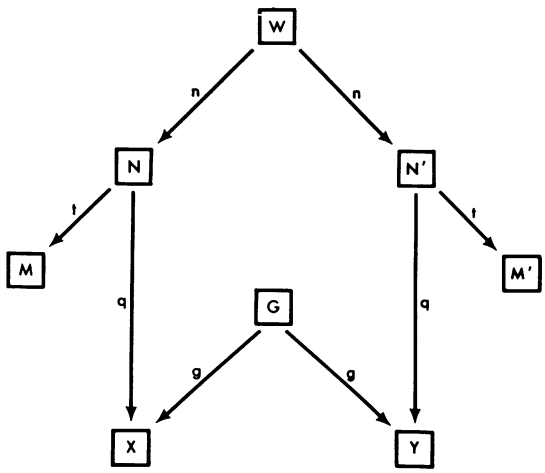

FIG. 7.—SSA. Full sibs reared apart by random foster parents. Common environments and their indices of the two foster homes, C and I for X and C' and I' for Y, are not shown, as in MZA.

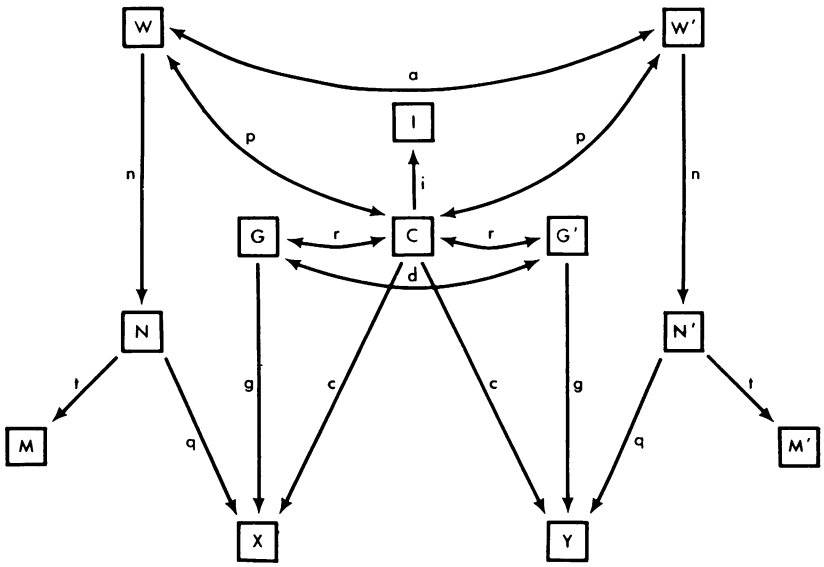

FIG. 8.—HST. Half-sibs reared together by their common parent. G and G' denote the midparent genotypes, and W and W' denote the midparent races for X and Y, respectively.

Half-Sibs Reared Together by Their Common Parent (HST) (Fig. 8)

The correlation between midparents with one parent in common is $d = (1 + 3m)/[2(1 + m)]$ for genotypes and $a = (1 + 3w)/[2(1 + w)]$ for race. It is assumed that the relevant environments of half-sibs reared together are no less alike than those of full sibs. Failure of this assumption (or appreciable dominance or epistasis) would tend to underestimate heritability. The half-sib correlation is

$$\rho_{XY} = c^2 + g^2d + 2grc + \{q^2n^2a + 2qnpc\}.$$

This situation is best realized when two alleged sibs are shown by paternity tests to be half-sibs.

Half-Sibs Reared Apart, One by True Parents, the Other by Random Foster Parents (HSP) (Fig. 9)

The critical assumptions are that the environments are uncorrelated (which is tested by $\rho_{II'} = 0$) and that the foster parents are not related to the true parents (which may be determined from the pedigrees). The half-sib correlation is

$$\rho_{XY} = (g^2d + \{q^2n^2a + qnpc\})\ \theta.$$

This relationship occurs when an extramarital child is placed for adoption at birth while a legitimate half-sib is retained.

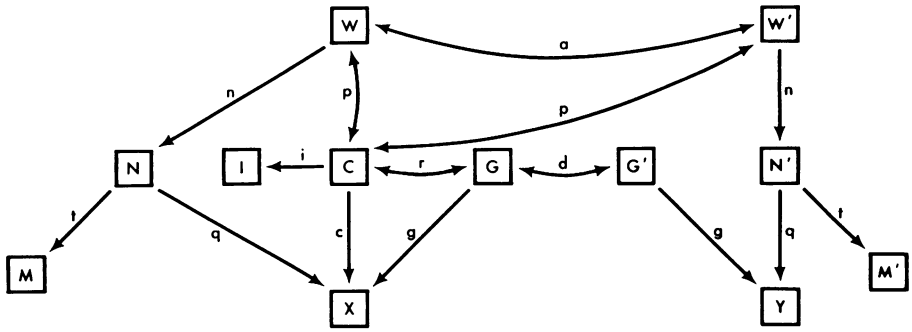

FIG. 9.—HSP. Half-sibs reared apart, X by true parents and Y by random foster parents. C' and I' for Y are not shown.

Half-Sibs Reared Apart by Random Foster Parents (HSA) (Fig. 10)

This relationship may be realized by an unmarried woman whose children by different fathers are placed for adoption at birth. The half-sib correlation is

$$\rho_{XY} = (g^2d + \{q^2n^2a\})\,\theta^2.$$

Half-Sibs Reared Apart by Their Own Parents (HSS) (Fig. 11)

This situation, which is common in some societies, is typified by concubinage or sequential monogamy. Considering the American divorce rate, it is surprising that this relationship has been neglected by human geneticists. It is most useful when the correlation b between home environments is negligible and, hence, $\rho_{II'} = 0$. The half-sib correlation is

$$\rho_{XY} = g^2d + c^2b + \{q^2n^2a\}.$$

Unrelated Foster Sibs (Adopted-Adopted) Reared Together by Random Foster Parents (FST) (Fig. 12)

The critical assumption is that differences in prenatal and postnatal environment before adoption had no phenotypic effect, which is difficult to test except

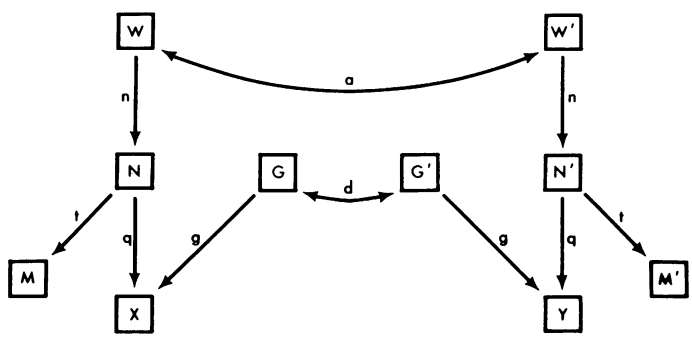

FIG. 10.—HSA. Half-sibs reared apart by random foster parents. C and I for X, as well as C' and I' for Y, are not shown.

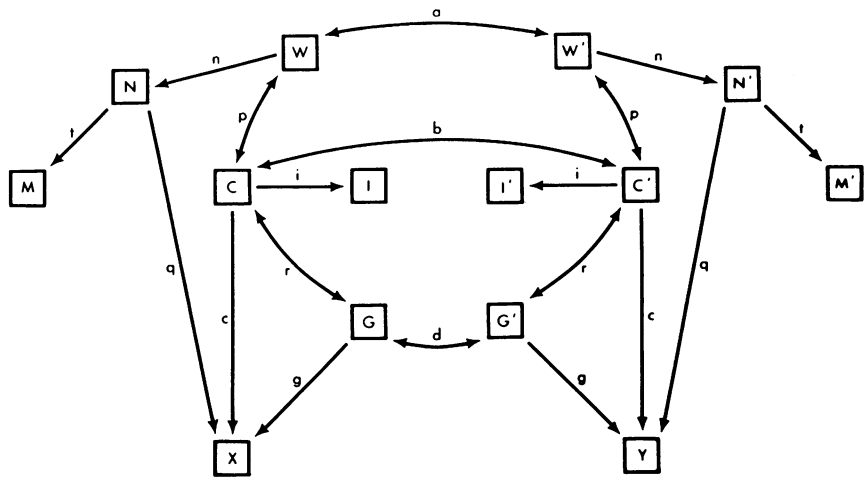

FIG. 11.—HSS. Half-sibs reared apart by their own parents

by overall goodness of fit. Violation of this assumption underestimates c and therefore inflates h. The foster sib correlation is

$$\rho_{XY} = c^2\theta^2.$$

Unrelated Foster Sibs (Natural-Adopted) Reared Together by the Parents of One of Them (FSP) (Fig. 13)

In this case the parents of a child subsequently decide to adopt an unrelated child, or they have a child of their own after first adopting an unrelated child. Differences in the prenatal and postnatal environments are assumed to have no phenotypic effect. The foster sib correlation is

$$\rho_{XY} = (c^2 + grc + \{qnpc\})\,\theta.$$

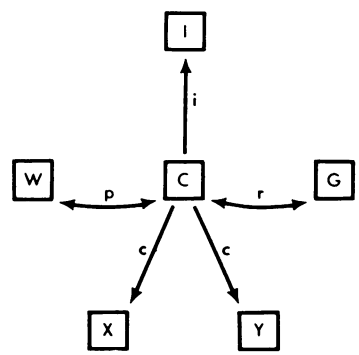

FIG. 12.—FST. Unrelated foster sibs (adopted-adopted) reared together by random foster parents. The race and its index, N and M of X and N' and M' of Y, are not shown.

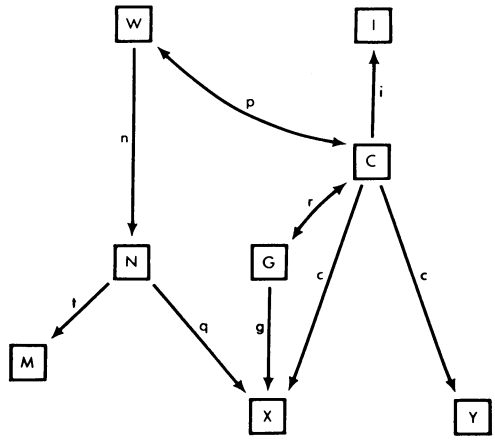

FIG. 13.—FSP. Unrelated foster sibs (natural-adopted) reared together by the parents of one of them (X). Race and index of Y, N' and M', are not shown.

Offspring-Parent, the Child Reared by His True Parent (OPT) (Fig. 14)

Familial correlations are more complicated between generations. We assume that the path from genotype to adult phenotype is hz and that the selective or modifying effect on the parent's adult phenotype of the home environment provided to his child is ck. On the null hypothesis of equal effects for children and adults, $z = k = 1$. The critical assumption is that parental phenotype influences the index only through C and that, by phenotypic selection or modification, C is a cause rather than an effect of X. Possibly these assumptions are too simple, but Wright [4] showed that more complicated ones lead to indeterminacy. Since our interest is in tests of hypotheses, the simplification does not seem objectionable;

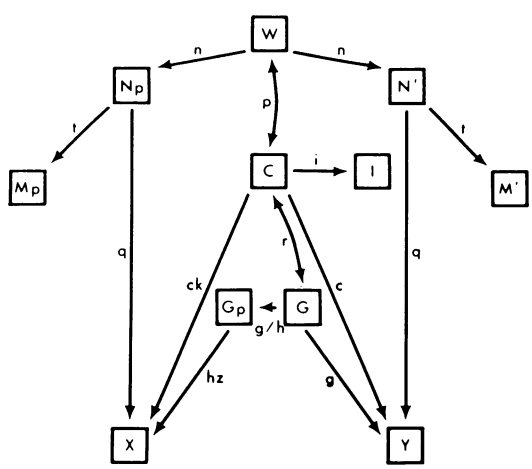

FIG. 14.—OPT. Offspring-parent, the child (Y) reared by his true parent (X). G_P, N_P, and M_P denote the genotype, race, and the racial index of the parent.

it may be avoided by considering only pairs in the same generation. The parent-offspring correlation is

$$\rho_{XY} = c^2k + g^2z + grc(z + k) + \{q^2n^2 + qnpc(1 + k)\}.$$

Offspring-Parent, the Child Reared by Random Foster Parents (OPA) (Fig. 15)

Like SSA, FST, and OFP, this is an important relationship for separating common environment from heritability. It is assumed that differences in prenatal and

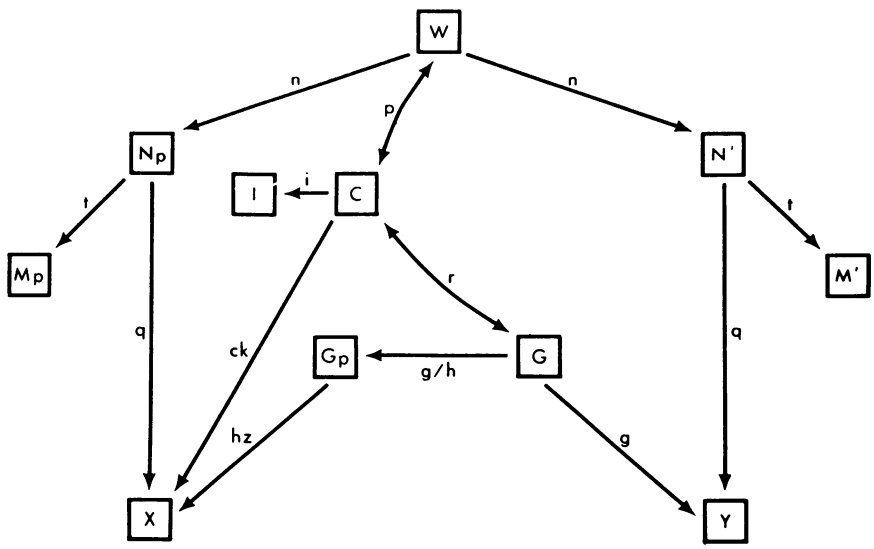

FIG. 15.—OPA. Offspring-parent, the child reared by random foster parents. C' and I' for Y are not shown.

postnatal environment before birth have no phenotypic effect and that the foster parents are random. Failure of these assumptions tends to inflate h. The correlation between offspring and true parent under these conditions is

$$\rho_{XY} = (g^2z + grck + \{q^2n^2 + qnpck\})\ \theta.$$

Offspring with Random Foster Parent (OFP) (Fig. 16)

The true parents are assumed to exercise no environmental effect on the child who is reared by random foster parents. The foster parent–child correlation is

$$\rho_{XY} = (c^2k + gzrc + \{qnpc\})\ \theta.$$

Like-sexed Twins (LS), Zygosity not Determined

We assume that the three types of twins have the following frequencies: opposite-sexed DZ, P; like-sexed DZ, P; and (like-sexed) MZ, $1 - 2P$. Among like-sexed twins, the proportion of dizygous pairs is $P/(1 - P)$. We may therefore

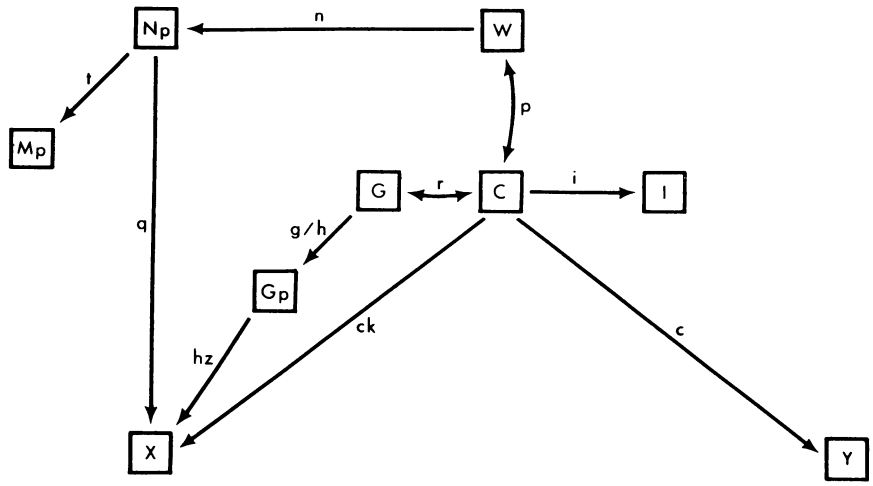

FIG. 16.—OFP. Offspring with random foster parent. N' and M' for Y are not shown.

write the expected correlation between two like-sexed twins (ρ_{LS}) as a function of expected correlations between monozygous twins (ρ_{MZ}) and dizygous twins (ρ_{SS}),

$$\rho_{LS} = \frac{P\rho_{SS} + (1 - 2P)\,\rho_{MZ}}{1 - P},$$

adjoining a third letter T, P, or A to LS, SS, and MZ to specify communality of the family environment. The estimate of P is taken to be without error. Since the analysis is both approximate and of reduced power, the investigator who wants to use twins would be well advised to determine their zygosity.

Other Equations

The enumeration of relationships may be extended, but only by introducing additional assumptions about transmission of common environment. It therefore seems reasonable to restrict attention to first- and second-degree biological or adoptive relatives, with emphasis on contemporaneous pairs.

For the various types discussed above, expected correlations among phenotype, index of common environment, and racial index may be easily derived (tables 4–6). Together they test hypotheses of familial correlation with greater power than alternative designs. Perhaps an understanding of these methods will encourage investigators to collect the necessary observations. Meanwhile, some instructive applications can be made to published data. Analysis of a given set of familial correlations is carried out by converting the correlations, both observed and expected, into z transforms and then applying the theory developed in the first section. All the necessary computations have been incorporated in a computer program, COMVAR (see Appendix).

TABLE 4

CORRELATIONS BETWEEN PHENOTYPE AND INDEX OF COMMON ENVIRONMENT
I = INDEX OF X; I' = INDEX OF Y; $\theta = 1/\sqrt{1 - 2(grc + qnpc)}$

VARIABLE	CORRELATION: General Term	+ Racial Term	RELATIONSHIPS FOR WHICH CORRELATION IS THE SAME
XI	$ci + gri$	$qnpi$	MZT, MZP, SST, SSP, HST, HSP, HSS, FSP
	$ci\theta$	0	MZA, MZF, SSA, HSA, FST
	$cki + gzri$	$qnpi$	OPT, OPA, OFP
XI'	$ci + gri$	$qnpi$	MZT, SST, HST, FSP
	$ci\theta$	0	MZF, FST
	cbi	0	HSS
	0	0	MZP, MZA, SSP, SSA, HSA, HSP, OPA
	$cki + gzri$	$qnpi$	OPT, OFP
YI	$ci + gri$	$qnpi$	MZT, SST, HST, OPT
	$gri\theta$	$qnpi\theta$	MZP, SSP, OPA
	0	$qnpi\theta$	HSP
	0	0	MZA, SSA, HSA
	cbi	0	HSS
	$ci\theta$	0	MZF, FST, FSP, OFP
YI'	$ci + gri$	$qnpi$	MZT, SST, HST, HSS, OPT
	$ci\theta$	0	MZP, MZA, MZF, SSP, SSA, HSP, HSA, FST, FSP, OPA, OFP

APPLICATIONS

Birth Weight

Studies by Robson [8] and Morton [9] have shown that "the resemblance in birth weight of sibs is largely attributable to the maternal constitution or environment, not to genetic similarity of sibs." Interracial crosses in Hawaii have birth weight intermediate between the incrosses, suggesting effects of the mother's postmarital environment [10].

The data of Morton [9] are summarized in table 7, where full sibs with one or more intervening births are omitted, since their resemblance is less than for twins or adjacent births. Table 8 shows that there is no significant difference between unlike-sexed twins and adjacent sibs and that the data cannot be explained by genetic heritability alone. However, a model of common maternal environment gives a good fit, which is not improved by introducing heritability ($\chi^2_1 = 0.48$). Assuming $h^2 = 0$, the estimate of c is .740 $\pm$ .018, corresponding to 55% of the phenotypic variance. The joint estimates are $\hat{h} = .290 \pm .207$ and $\hat{c} = .703 \pm .058$, which account for 8% and 49%, respectively, of the phenotypic variance. We see that the estimate of maternal effect and its significance do not depend critically on whether heritability is assumed null or estimated simultaneously. Conversely, there is little information about heritability when the effect of common

TABLE 5

CORRELATIONS BETWEEN PHENOTYPE AND RACIAL INDEX
M = RACIAL INDEX FOR X; M' = RACIAL INDEX FOR Y; $\theta = 1/\sqrt{1 - 2(grc + qnpc)}$

Variable and Correlation	Relationships for Which Correlation is the Same
XM:	
$qt + cpnt$	MZT, MZP, SST, SSP, HST, HSP, HSS, FSP
$qt\theta$	MZA, MZF, SSA, HSA, FST
$qt + ckpnt$	OPT, OPA, OFP
XM':	
$qt + cpnt$	MZT, MZP, SST
$qt\theta$	MZA, MZF
$qn^2t + cpnt$	SSP
$qn^2t\theta$	SSA
$qn^2at + cpnt$	HST, HSP
$qn^2at\theta$	HSA
qn^2at	HSS
$qn^2t + ckpnt$	OPT, OPA
0	FST, FSP, OFP
YM:	
$qt + cpnt$	MZT
$qt\theta$	MZP, MZA, MZF
$qn^2t\theta$	SSP, SSA, OPA
$qn^2at + cpnt$	HST
$qn^2at\theta$	HSP, HSA
qn^2at	HSS
$cpnt\theta$	FSP, OFP
$qn^2t + cpnt$	SST, OPT
0	FST
YM':	
$qt + cpnt$	MZT, SST, HST, HSS, OPT
$qt\theta$	MZP, MZA, MZF, SSP, SSA, HSA, HSP, FST, FSP, OPA, OFP

environment is important. For the joint estimate, the element of the information matrix for h is $K_{hh} = 205$, whereas the reciprocal of the corresponding element of the inverse (covariance) matrix is only $1/K^{hh} = 23$. The loss of information about h when c must be estimated simultaneously is $1 - 1/K_{hh}K^{hh} = 89\%$. The importance of a maternal effect is confirmed, but neither the role of the mother's genotype nor the values of h, m, or r can be determined from these data.

Race, Social Class, and IQ

Scarr-Salapatek [11] reported aptitude tests on 992 twin pairs in Philadelphia schools. An additional 124 pairs were omitted because one or both members were enrolled in special classes. Selection must have been more severe for lower class blacks (mean score 27.7) than for higher class whites (mean 50.9). Zygosity was not determined, but the composition of like-sexed twins was estimated. Pairs were classified as of black or white race and below or above the median social class. Intraclass correlations were reported for each group separately for verbal,

TABLE 6

CORRELATIONS BETWEEN RACIAL INDEX AND INDEX OF COMMON ENVIRONMENT
M, I = INDICES OF X; M', I' = INDICES OF Y

Variables and Correlation	Relationships for Which Correlation Is the Same
MI:	
tnpi	MZT, MZP, SST, SSP, HST, HSP, HSS, FSP, OPT, OPA, OFP
0	MZA, MZF, SSA, HSA, FST
MI':	
tnpi	MZT, SST, HST, FSP, OPT, OFP
0	MZP, MZA, MZF, SSP, SSA, HSS, HSP, HSA, FST, OPA
M'I:	
tnpi	MZT, MZP, SST, SSP, HST, HSP, OPT, OPA
0	MZA, MZF, SSA, HSS, HSA, FST, FSP, OFP
M'I':	
tnpi	MZT, SST, HST, HSS, OPT
0	MZP, MZA, MZF, SSP, SSA, HSP, HSA, FST, FSP, OPA, OFP
II':	
i^2b	HSS
1	MZT, MZF, SST, HST, FST, FSP, OPT, OFP
0	MZP, MZA, SSP, SSA, HSA, HSP, OPA
MM':	
t^2n^2	SST, SSP, SSA, OPT, OPA
t^2n^2a	HST, HSP, HSS, HSA
1	MZT, MZP, MZA, MZF
0	FST, FSP, OFP

TABLE 7

HUMAN BIRTH WEIGHT (MORTON [9])

Source	Code	No. Pairs	Intraclass Correlation
Like-sexed twins	LST	220	.557
Unlike-sexed twins	SST	40	.655
Adjacent sibs	SST	367	.523
Maternal half-sibs	HST	30	.581
Paternal half-sibs	HSA	168	.102

TABLE 8

ANALYSIS OF HUMAN BIRTH WEIGHT CORRELATIONS

Source	df	χ^2
Between unlike-sexed twins and adjacent sibs	1	1.61
$\hat{h}$: $m=c=r=0$	3	33.75
$\hat{c}$: $m=h=r=0$	3	2.09
$\hat{h},\hat{c}$: $m=r=0$	2	1.61

nonverbal, and total IQ scores. Heritability for the ith group was estimated by $h_i^2 = 2(r_{iMZ} - r_{iDZ})$, where r_{iDZ} was the observed intraclass correlation for unlike-sexed twins and r_{iMZ} was an estimate of the corresponding value for monozygous twins. Without reporting any significance tests, she concluded that "population differences in heritability of IQ scores were found for racial and social class groups."

Eaves and Jinks [12] examined her data and decided that "evidence previously analyzed is insufficient to support the conclusions drawn." They based this conclusion on z transforms of the correlations for the eight groups defined by race, class, and concordance for sex. Heterogeneity was not found among the eight correlations for nonverbal scores ($\chi^2_7 = 5.63$), but verbal scores were barely heterogeneous ($\chi^2_7 = 15.63$). However, an unweighted factorial design showed that none of the three main effects (race, class, and sex) or interactions was significant. Therefore,

> the only tenable conclusion to be drawn from the data is that there is a highly significant correlation between twins of all kinds for verbal IQ. . . .
>
> We are in no position to decide the cause of such similarity. There is no evidence that it has a genetical basis as far as this study goes, but as we have shown above, the likelihood of detecting such an effect with this experimental design and with these samples is very small. There is certainly no evidence in Scarr-Salapatek's studies that the proportion of genetical variation in either verbal or non-verbal IQ depends on race or social class. In view of this conclusion, and having regard to the general absence of genotype-environmental interactions for IQ, there is little justification for detailed consideration of the particular models suggested by Dr. Scarr-Salapatek.

We reanalyzed the data by our methods (table 9). Bartlett's test for homogeneity of variances gives no suggestion that the variance of individuals is heterogeneous among groups. Evidently, sample sizes are too small to detect differential effects of excluding special-class students. The hypothesis that a deprived environment reduces aptitude variance is not supported.

Analysis of z scores is in agreement with Eaves and Jinks. The small discrepancy between their χ^2s and ours is trivial: they may have omitted the bias correction in z or the small-sample correction in σ_z^2. Partition of the heterogeneity χ^2 shows that there is significant variation between race-class groups for verbal and total aptitude, with no significant heterogeneity between like-sexed and unlike-sexed twins within race-class groups. By this method, race and class effects are not separated: there may be racial or class variation, or both.

The data are elucidated by fitting a linear model. Since there are only two equations, for like-sexed and unlike-sexed twins, only two parameters can be considered. We took common environment (c) and genotype (h), assuming $r = m = 0$. The special case $h = 0$ gives $\hat{c} = .792$ for total aptitude (combining all race-class groups), with a nonsignificant residual χ^2. When $\hat{c}_i$ is determined for each group, conditional on either $h = 0$ or the overall estimate of h, the residual

TABLE 9

VERBAL, NONVERBAL, AND TOTAL APTITUDE SCORES OF TWINS (SCARR-SALAPATEK [11])

SOURCE	df	χ^2		
		Verbal	Nonverbal	Total Aptitude
Homogeneity of total variances	7	5.25	3.11	8.84
Analysis of correlation coefficients:				
Among eight correlations	7	13.40	5.77	14.99
Within races	6	8.55	2.11	5.53
Between races	1	4.85	3.66	9.46
Within classes	6	5.76	4.08	8.08
Between classes	1	7.64	1.69	6.91
Within race-class groups	4	4.35	1.54	2.99
Between race-class groups	3	9.05	4.23	12.00
Fitting models:				
$\hat{c}$: $h = 0$ (overall)	1	3.22	0.18	0.16
$\hat{c}_i$: $h_i = 0$	4	4.35	1.54	2.99
$\hat{c}_i$: $h_i = h$	3	1.12	1.46	2.64
$\hat{h}$: $c = 0$ (overall)	1	3.29	10.51	16.78
$\hat{h}_i$: $c_i = 0$	4	4.39	12.91	18.37
$\hat{h}_i$: $c_i = c$	3	0.75	1.68	1.75

χ^2 is nonsignificant. Thus, there is no evidence of heritability, and any variation among groups may be limited to the c_i. The results are quite different for models which assume that $c = 0$. Overall, $c = 0$ gives a poor fit for total and nonverbal aptitude and a marginal fit for verbal aptitude: common environment cannot be neglected. There is significantly poor fit for nonverbal and total aptitude when c_i is assumed zero, but this disappears when the overall estimate of c is used for each c_i.

The conclusions from this analysis are simple. With zygosity undetermined and twins reared together, the sample sizes are too small to determine heritability, let alone to establish whether it varies among race or class groups. However, there is some kind of heterogeneity in twin correlations, most economically attributed to common environment, which could only be explained by a study of larger and better design. Scarr-Salapatek was right in suspecting (without a significance test) that the groups were heterogeneous. Eaves and Jinks were correct in concluding that heritability differences are unproven; their failure to detect race or class effects was due to use of a complete factorial design on samples too small to separate race from class or heritability from common environment. Their nonsignificant results correspond to $\hat{c}_i$: $h_i = h$ and $\hat{h}_i$: $c_i = c$ of table 9, but they missed the significant effects revealed by a hierarchical analysis.

Tender-mindedness

Cattell et al. [13] analyzed small samples of monozygous and dizygous twins, sibs, and foster children reared together, and sibs reared apart. Their table 1 gives

intrapair variances, from which intraclass correlations may be obtained as the complement of the ratio of intrapair variance to the variance for their large sample of unrelated reared apart. Twelve "factors" of personality were considered, the first of which is tender-mindedness, for which they concluded: "this pattern is largely environmentally determined, and on more accurate analysis might prove to be wholly an environmental mold trait. The larger ratio for *between* family environment sugggests that this resides in some sort of family atmosphere—almost certainly an over-protective, gentler tradition as opposed to spartan roughness. The size of positive correlation suggests some selection of gentler temperaments to the gentler environment."

This interpretation seems plausible if tender-mindedness is measured by factor 1, but the data provide little support. Table 10, which incorporates the analysis of

TABLE 10

TENDER-MINDEDNESS (CATTELL ET AL. [13])

Hypothesis	df	χ^2
Uncorrected variances:		
$\hat{h}$: $c=m=r=0$	3	4.14
$\hat{c}$: $h=m=r=0$	3	4.45
$\hat{h},\hat{c}$: $m=r=0$	2	0.79
Corrected variances:		
$\hat{h}$: $c=m=r=0$	2	3.86
$\hat{c}$: $h=m=r=0$	2	4.31
$\hat{h},\hat{c}$: $m=r=0$	1	0.53

both "uncorrected" and "corrected" variances for factor 1, shows that neither heritability nor common environment has a significant effect, with allowance for the other, nor does gene-environment correlation approach significance; following [13], we dropped full sibs reared apart from the analysis of corrected variances. The suggestion of Cattell et al. of strictly environmental determination is at variance with their unconfirmed claim of association between factor 1 and the A blood group [14], which has been criticized on statistical grounds [15].

On the evidence, nothing much can be said about the roles of nature and nurture for tender-mindedness and other personality traits.

IQ of True and Adopted Children

Since the classical study of Burks [16], many investigations of familial resemblance for IQ have been made [17]. Relevant American data are summarized in table 11; they are from a restricted range of families and may not apply to a particular social class. The phenotype is Binet IQ for children and MA (mental age) for adults. The index of common environment is Burks's culture index.

The three studies of natural-adopted pairs (FSP) reported by Jencks [17,

TABLE 11

IQ OF TRUE AND FOSTER CHILDREN (BURKS [16] AND JENCKS [17])

Relation	Sample Size	Correlation
Burks [16]:		
FSTXY	21	.23
SSTXI (culture index)	101	.44
SSAXI	186	.25
OPTXI	100	.67
OPTXI	105	.71
Jencks [17]:		
OPTXY (Jencks's table A-2)	200	.46
	366	.51
	441	.49
	102	.58
	141	.35
OFPXY (X = father) (Jencks's table A-3)	178	.07
	180	.37
	178	.19
OFPXY (X = mother)	204	.19
	255	.28
	186	.24
FSPXY (Jencks's table A-9 [pooled])	94	.26
SSTXY	1,951	.52
SSTXY (DZ)	50	.63
MZTXY	50	.89
MZAXY (Jencks's table A-12 [U.S.; Binet])	19	.69

table A-8] are homogeneous ($\chi^2_2 = 1.81$), but the correlations for adopted-adopted pairs (FST) are significantly heterogeneous ($\chi^2_4 = 9.79$), because of the large value of .65 reported by Skodak [18]. Jencks discusses the anomalous result of pooling these studies, with the correlation for adopted-adopted pairs exceeding natural-adopted pairs, and in effect decides to discount Skodak's result. We have chosen instead to use only Burks's value for adopted-adopted pairs (.23), which as expected is slightly less than the pooled value of .26 for natural-adopted pairs. Both are in agreement with English studies which give .25 for all unrelated children reared together [17, table A-9]. Other correlations are not significantly heterogeneous within any type of relationship (table 12, $\chi^2_{11} = 17.88$).

The data of table 11 are analyzed in table 12, assuming all correlations are interclass; essentially the same results are obtained even if the phenotypic correlations are treated as intraclass. The complete model with seven parameters (h, m, c, r, k, z, i) gives a good fit; the estimate of m is $-.054 \pm .111$, and the hypothesis that $m = 0$ cannot be rejected ($\chi^2_1 = 2.80 - 2.24 = 0.56$). Presumably m is positive, but a large value can be excluded: for example, $m = .2$ gives a significantly poor fit ($\chi^2_1 = 8.79 - 2.24 = 6.55$). Apparently the observed correlation between husbands' and wives' test scores, which is estimated in table A-1

TABLE 12

ANALYSIS OF IQ DATA OF TABLE 11

Source	df	χ^2
Between OPTXY correlations	4	6.10
Between OFPXY correlations	5	10.32
Between SSTXY correlations	1	1.17
Between OPTXI correlations	1	0.29
Pooled	11	17.88
Testing hypotheses:		
$h = m = r = 0,\ k = z = 1$	8	186.68
$m = r = 0,\ k = z = i = 1$	8	41.40
$m = r = 0,\ k = i = 1$	7	36.90
$m = r = 0,\ z = i = 1$	7	34.02
$m = 0,\ k = z = 1$	6	22.55
$m = r = 0$	5	6.85
$m = 0$	4	2.80
$m = .2$	4	8.79
$r = 0$	4	6.02
fitting h, m, c, r, k, z, i	3	2.24
$m = r = 0,\ i = 1$	6	6.85
$m = 0,\ k = z = i = 1$	7	22.57
$m = 0,\ i = 1$	5	3.50
$r = 0,\ i = 1$	5	6.03
$i = 1$	4	3.36

of Jencks [17] as .50, is not due to the genetic correlation m. The study of Higgins et al. [19], which used school records, gave .33. This low value may be due to diversity of records, but it suggests an effect of postmarital environment. Taking $m = 0$, maximum-likelihood estimates of the other six parameters were obtained (table 13). Table 14 presents analysis of z transformations at these parameter values.

The correlation between Burks's culture index and common environment is estimated for $m = 0$ as $\hat{i} = .902 \pm .119$, which is not significantly less than 1 ($\chi^2_1 = 1.12$). This justifies the assumption that $i = 1$ in the model of Wright [4].

TABLE 13

ESTIMATES OF PATH COEFFICIENTS FOR IQ DATA OF TABLE 11 (WITH $m = 0$)

Path Coefficient	Estimate	Standard Error
h	0.822	0.035
c	0.286	0.044
r	0.326	0.195
k	2.402	0.691
z	0.402	0.242
i	0.902	0.119

TABLE 14

ANALYSIS OF z TRANSFORMS FOR ESTIMATES OF TABLE 13

RELATION	z TRANSFORM		INFORMATION	
	Observed	Expected	$1/\sigma_z^2$	χ^2
MZTXY	1.413	1.318	47.2	0.42
MZAXY	0.829	0.993	16.5	0.44
SSTXY	0.580	0.587	1,996.4	0.11
FSTXY	0.229	0.092	18.9	0.35
FSPXY	0.265	0.145	91.9	1.32
OPTXY	0.528	0.528	1,238.8	0
OFPXY	0.232	0.235	1,168.6	0.01
SSTXI	0.470	0.458	98.8	0.01
SSAXI	0.255	0.280	183.9	0.12
OPTXI	0.846	0.844	200.0	0.00

NOTE.—Total χ^2 (df $= 10 - 6 = 4$) $= 2.80$.

However, complete determination of I by C seems unlikely, so we have retained the estimate of i.

Gene-environment correlation may be appreciable ($r = .326 \pm .195$) but is barely significant ($\chi^2_1 = 6.85 - 2.80 = 4.05$). This provides modest support of the hypothesis that class differences in mean IQ have a genetic component. Good designs and enormous samples would be necessary (and might not be sufficient) to establish with precision a genetic component in class (or race) differences.

There is a striking shift from childhood to maturity in the effects of nature and nurture ($k = 2.4$, $z = 0.4$). The hypothesis that $k = z = 1$ is untenable ($\chi^2_2 = 19.75$). This was noticed by Wright [20], who remarked, "It turned out to be mathematically impossible to assign the same values to the path coefficients of the parental generation as in the offspring generation, but this is not surprising since the parents were tested as adults instead of young children." In the parental generation he reported coefficients to the nearest .05 because "the solution is not strictly determinate, but is so within rather narrow limits." Table 15 shows that Wright's estimates of variance components are in at least qualitative agreement with ours. Assuming an attenuation factor of $\alpha = .1$, genotype is estimated to account for most of the true variance of IQ in children but only a minor fraction of adult variance. Conversely, the effect of common environment increases. This is all the more remarkable because common environment refers to the conditions under which the adult rears his child. As Wright noted, "Midparent IQ is a much better index of home environment than of child's heredity." It is a mark of Wright's insight that subsequent investigations which make the path diagram determinate have not substantially altered his conclusions.

Jencks made a number of quantitative assumptions and failed to distinguish juvenile and adult phenotypes in his path diagrams. Nevertheless, his estimates (table 15) are not far from the means of the two generations, with some inflation of gene-environment covariance at the expense of environment.

TABLE 15

ESTIMATES OF VARIANCE COMPONENTS (AS FRACTIONS OF TOTAL VARIANCE) FOR IQ DATA OF TABLE 11, ASSUMING $\alpha = .1$ AND $m = 0$ (ESTIMATES OF TABLE 13 ARE USED HERE)

SOURCE	CHILDREN		PARENTS		GENERATIONS POOLED (JENCKS [17])
	Our Analysis	Wright [4]	Our Analysis	Wright [4]	
Genotype	.752 ± .058	.50	.121 ± .152	.30	.45
Genotype-environment covariance	.120 ± .054	.12	.117 ± .032	.14	.20
Common environment	.091 ± .025	.07	.523 ± .178	.36	.35
Random environment	.037 ± .032	.31	.239 ± .089	.20	
Total	1	1	1	1	1

DISCUSSION

The above examples show that better designs and much larger samples are required to discriminate hypotheses than to estimate effects under a particular hypothesis. Foster children or half-sibs are essential for separating nature and nurture if common environment is important but are not sufficient if common uterine environment is important (as may perhaps be true for monozygous twins). In sufficiently large samples such discrepancies should be detected by significant deviations from our model. This conclusion presumably holds for racial differences as well.

We have assumed that placement of foster children is random, which is the most favorable situation for genetic analysis. However, it is not difficult to extend our models to assortative adoption, introducing a correlation between common environments of the true and foster homes (similar to b for half-sibs reared by their own parents). The correlation between indices, $r_{II'}$, is informative; a positive value would suggest assortative adoption.

Our treatment of parent-offspring pairs has been simplified, since most studies of familial correlation deal with children. If our scheme is not adequate, environment of the parent as a child may be introduced, connected to the environment he provides his children by a direct path and a path through his adult phenotype. Relations between adult siblings and between indices of common environments in successive generations would be useful in determining the values of these paths. However, the large standard errors for parents in table 15 suggest that enormous samples would be required to benefit from this sophistication.

Complex segregation analysis of quantitative traits is a powerful method for resolving major genes and polygenes but neglects environment common to parents and children [21]. Recently we introduced dominance into path analysis to provide a test in the presence of common environment. We do not yet have experience

with this extension, which will be reported in a later paper. In principle, path analysis can discriminate dominance from common environment, whereas segregation analysis separates a major locus with dominance from additive polygenes. Only by a combination of the two approaches can all the mechanisms (common environment, dominance, polygenes) be resolved.

Significance tests in path analysis are like any new method: the limitations will be minimized by some and exaggerated by others. We have favored correlations because they provide excellent tests under z transformation. This makes our results sensitive to selection biases, especially those which affect the phenotypic variance and so can easily be detected. Path regressions are more robust but are still sensitive to phenotypic selection, which in practice is more common than selection on independent variables. It seems to us that there are enough problems in human biometrical genetics without introducing biased selection, which is sufficiently protean to invalidate any path analysis, for which we regard random sampling as an essential condition.

Another limitation is to linear systems, which seems to us justified on several grounds: (1) no practical method for analysis of nonlinear systems has been developed; (2) unless interaction produces a significant discrepancy from a simple linear model, we must conclude that it is negligible at the observed sample sizes; (3) even intense epistasis typically produces little nonadditive variance [22]; (4) it would be important to separate gene-environment interaction from the environmental effect only in the unlikely event that cloning becomes feasible and widely practiced; and (5) we are much less interested in estimating a component of variability (like h^2) than in testing a null hypothesis. Since armchair examples of significant interaction in the absence of an additive effect are pathological and have never been demonstrated in real populations, we need not be unduly concerned about interaction effects. The investigator with a different view should publish any worthwhile results he may obtain.

SUMMARY

From Wright's path analysis, a theory of hypothesis testing for linear models of familial correlation is developed and applied to several bodies of data, for some of which the authors' conclusions are not supported. Much larger samples and better designs are required to discriminate hypotheses than to estimate parameters under a particular model. Possible extensions of the method are discussed.

APPENDIX

COMVAR, a computer program for the analysis of components of variation in families, is written in FORTRAN IV for the CDC 3100 computer. Each data card gives information about one correlation, indicating the relationship and the variates (MZTXY, SSTXI, etc.), value of the correlation, number of pairs of observations on which it is based, and whether it is interclass or intraclass. A preliminary analysis pools correlations for the same type of relationship and the same variates and gives a χ^2 test of homogeneity.

For the main analysis, one control card gives initial values of the parameters, with a zero default option. Another control card specifies which parameters are to be estimated

by Newton-Raphson iteration. These two control cards may be repeated for a given set of data. Iterations continue until all the correction factors become less than 5×10^{-6} in absolute value, or the information matrix becomes singular, or 50 iterations are made. Estimates are confined to the possible range of the parameters (between -1 and $+1$ for correlations, between 0 and $+1$ for other paths). For the first and last iterations, a table is printed giving observed and expected values of z, information, goodness-of-fit χ^2 (as in table 14), and $u'K^{-1}u$, the quadratic form testing convergence of the estimates. Upon convergence, final estimates, standard errors, u scores, the information matrix, and the covariance matrix are printed. An optional control card giving the value of α (attenuation correction factor) prints components of variance with their large-sample standard errors. for both children and parents.

All analyses presented in this paper were done with COMVAR, which is part of the program library of the Population Genetics Laboratory and the World Health Organization International Reference Center for Processing of Human Genetics Data, from which copies of the program and a detailed description are available.

REFERENCES

1. WRIGHT S: Correlation and causation. *J Agric Res* 20:557–585, 1921
2. WRIGHT S: *Genetics and Biometric Foundations,* vol. 1. Chicago, Univ. Chicago Press, 1968
3. WRIGHT S: The origin of the *F*-statistics for describing the genetic aspects of population structure, in *Genetic Structure of Populations,* edited by MORTON NE, Honolulu, Univ. Press Hawaii, 1973
4. WRIGHT S: Statistical methods in biology. *J Am Stat Assoc* 26:155–163, 1931
5. MORTON NE: Analysis of family resemblance. I. Introduction. *Am J Hum Genet* 26: 318–330, 1974
6. KENDALL MG, STUART A: *The Advanced Theory of Statistics,* vol. 2. New York, Hafner, 1961
7. FISHER RA: On the probable error of a coefficient of correlation deduced from a small sample. *Metron* 1:1–32, 1921
8. ROBSON EB: Birth weight in cousins. *Ann Hum Genet* 19:262–268, 1955
9. MORTON NE: The inheritance of human birth weight. *Ann Hum Genet* 20:125–134, 1955
10. MORTON NE, CHUNG CS, MI MP: Genetics of interracial crosses in Hawaii. *Monographs in Human Genetics,* vol. 3. Basel, Karger, 1967
11. SCARR-SALAPATEK S: Race, social class, and IQ. *Science* 174:1285–1295, 1971
12. EAVES LJ, JINKS JL: Insignificance of evidence for differences in heritability of IQ between races and social classes. *Nature* (Lond) 240:84–88, 1972
13. CATTELL RB, BLEWETT DB, BELOFF JR: The inheritance of personality: a multiple variance analysis determination of approximate nature-nurture ratios for primary personality factors in *Q*-data. *Am J Hum Genet* 7:122–146, 1955
14. CATTELL RB, YOUNG HB, HUNDLEBY JD: Blood groups and personality traits. *Am J Hum Genet* 16:397–402, 1964
15. NORTON HW: Blood groups and personality traits. *Am J Hum Genet* 23:225, 1971
16. BURKS BS: The relative influence of nature and nurture upon mental development: a comparative study of foster parent–foster child resemblance and true parent–true child resemblance, in *27th Yearbook of the National Society for the Study of Education,* pt. 1. Bloomington, Ind., Public School Publishing Co., 1928, pp 219–316
17. JENCKS C: *Inequality.* New York, Basic Books, 1972
18. SKODAK M: Mental growth of adopted children in the same family. *Pedogogical Seminary and J Genet Psychol* 77:3–9, 1950

19. HIGGINS JV, REED EW, REED SC: Intelligence and family size: a paradox resolved. *Eugen Q* 9:84–90, 1962
20. WRIGHT S: The method of path coefficients. *Ann Math Stat* 5:161–215, 1934
21. MORTON NE, MACLEAN CJ: Analysis of family resemblance. III. Complex segregation analysis of quantitative traits. *Am J Hum Genet.* In press, 1974
22. CROW JF, KIMURA M: *An Introduction to Population Genetics Theory.* New York, Harper & Row, 1970, pp 124–129