Genetic and Cultural Inheritance of Continuous Traits

SARAH P. OTTO

Department of Integrative Biology
UC Berkeley, Berkeley CA 94720

FREDDY B. CHRISTIANSEN

Department of Ecology and Genetics
University of Aarhus
DK-800 Aarhus C, Denmark

MARCUS W. FELDMAN

Department of Biological Sciences
Stanford, CA 94305

July 28, 1994

Introduction

Mathematical and/or verbal models are intended to abstract the main features of causative relationships that exist in nature. Uncertainty as to whether any such model accurately describes these relationships or is merely consistent with observed data is particularly an issue in analyses of human behavior. A person's score on a behavioral test may be influenced in complex ways by family, friends, health, weather, and an unknown number of other factors. Further, the basic properties of an individual's reactions may change over time and other individuals may respond to these changes in behavior. Complicated interactions between the effects of different experiences, some complimentary and some antagonistic, may occur during any phase of development. Observations on variation in human behavior can therefore provide only a very incomplete description of causal relationships that are too complex to be completely specified by any model.

Path analysis represents an attempt to describe the etiology of complex traits under the assumption that all influences are linear without higher-order interactions, which makes it possible to directly compute expected correlations between relatives. Thus the entire process of inheritance and experience is modeled as a set of linear relationships between variables in individuals and their parents. Usually, the phenotype of an individual is viewed as a linear combination of its *genotypic value* (measured on the same scale as the phenotype) transmitted from its mother and father by some rule intended to approximate known rules of genetic transmission, its *cultural value* (measured on the phenotypic scale) which may be transmitted directly or indirectly from parents, and a non-transmitted *environmental value* (again measured on the phenotypic scale) whose rules of transmission do not involve the parents but may be thought of as summarizing those experiences of the individual that have influenced its phenotype.

In this paper, we review different linear models that specify how the phenotype of an individual is determined and how genetic and cultural effects are transmitted between generations. In the first section we describe in detail how an individual's phenotype might be influenced by its genes, cultural experiences, and non-heritable environmental experiences. The transmission of genetic effects from parents to offspring assumes a completely additive genetic model and Mendelian inheritance. While genetic transmission follows biometric models based on Mendel's rules, a similar frame

of reference is not available for models of cultural inheritance. We will assume, however, that cultural transmission occurs vertically; that is, cultural attributes are transmitted only from parents to offspring. We focus on models that allow genetic and cultural effects to pass from generation to generation with both parents contributing equally, but this unbiased vertical transmission is clearly only one of several possible mechanisms of cultural transmission (Cavalli-Sforza and Feldman, 1981). We consider two different cultural-inheritance models, an indirect-inheritance model where latent cultural attributes are transmitted and a direct-inheritance model where the observable phenotypes of parents directly influence the cultural attributes of a child.

Years of research on heritable variation in human populations have shown the importance of deviations from random mating, especially assortative mating. Different phenomena may cause assortment and we discuss two: *phenotypic homogamy,* where mates are chosen on the basis of phenotypic similarities, and *social homogamy,* where mates are chosen according to common membership in a social group.

Specification of the factors that influence the phenotype and a model of their inheritance allow the derivation of expected correlations between relatives, biological relatives as well as relatives whose only ties are cultural. We concentrate in the text on the description of relations among primary relatives and defer correlations between distant relatives to an appendix. These correlations are compared to observed correlations in different published sets of data for human behavioral characters (IQ and personality measures). Only three of the many possible linear models are used in this coarse statistical analysis, but they may serve as a basis for discussions of the limitations inherent in the model construction.

The models and the data used here are both limited in their ability to describe human behavior. Rather gross simplifications are necessary to keep the structure described in the models comparable to the limited structure displayed by the data. The models all assume multifactorial inheritance with genetic additivity in which such phenomena as dominance and epistasis are ignored. Also, the studied populations are assumed to be at equilibrium, i. e., neither the variances nor the means for any of the variables are allowed to change over time. The data are correlations gathered from human populations, so that the application of experimental controls employed by plant and animal breeders in their studies of inheritance is precluded.

The interpretation of a set of data applies to the variation in human behavior

within the population studied, at the time it was sampled. The data are taken from individuals who have been exposed to a specific set of environments, and analysis of these data cannot suggest how changing the developmental milieu might have altered the phenotype under study. Therefore, the results cannot tell us how to change environmental factors to change a person's intelligence, happiness or friendliness. At best we can obtain a crude indication of the importance of factors describing genetic and cultural inheritance for the *observed variation* in the population using models that make extremely simple assumptions about behavior. As we shall see, even these simple models differ in their estimates of the factors important in trait variation. The models that we discuss are those commonly used to produce estimates of heritability, and we shall see that these heritability estimates are indeed sensitive to the model assumptions even within the context of simple linear biometric models.

1 Transmission Models

The phenotype of an individual will be specified by a linear combination of genotypic, cultural, and environmental contributions. As usual in quantitative genetics, the phenotypic value of an individual, denoted by the variable P, is represented as a deviation from the mean in the population so that its expectation may be set to zero, $\mathrm{E}(P) = 0$. Also, the variance of the phenotypic variable is normalized to unity, $\mathrm{Var}(P) = 1$, which amounts to a simple change in the unit of measurement for the phenotype. The effect of the genotype of the individual on its phenotype is denoted by the variable A, with $\mathrm{E}(A) = 0$ by definition. The environmental contribution to an individual's phenotype will be divided into two parts; an effect influenced by cultural transmission, B, and a non-transmitted environmental effect, E, dependent only on the particular environmental experiences of the individual, with $\mathrm{E}(B) = \mathrm{E}(E) = 0$. All variables are assumed to be normally distributed and any collection of variables follows a multidimensional Gaussian distribution. In addition, the variances of the variables are all normalized to unity, $\mathrm{Var}(A) = \mathrm{Var}(B) = \mathrm{Var}(E) = 1$, and we use the parameters, h, b and e, to describe the strength of the influence of genes, cultural environment, and non-transmitted environment, respectively, on the phenotype of an individual. The phenotype of an individual is then specified as a linear combination

of the normalized genetic and environmental deviations:

$$P = hA + bB + eE. \tag{1}$$

The genetic (A) and environmental (B and E) deviations (the "hidden" variables) are measured in such a way that the parameters, h, b and e are positive and each is bounded by 0 and 1. In general, a covariance w will exist between the genetic deviation and the cultural deviation of the individual, $w = \mathrm{Cov}(A, B)$, but we assume that all transmissible components are independent of the non-transmitted environmental component, i. e. $\mathrm{Cov}(E, A) = \mathrm{Cov}(E, B) = 0$. The phenotypic variance is therefore

$$\mathrm{Var}(P) = h^2 + b^2 + e^2 + 2hbw, \tag{2}$$

and, since $\mathrm{Var}(P) = 1$, the strength of the non-transmitted environmental effects (e) must satisfy:

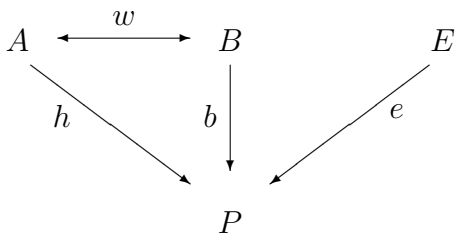
$$e = \sqrt{1 - h^2 - b^2 - 2hbw}. \tag{3}$$

Equation (1) describes the phenotypic deviation of an individual from the population mean, measured in standard deviates. This relative deviation of the phenotype is given in terms of a relative genetic deviation, a relative cultural deviation and a relative environmental deviation. The principal parameters of the model, h and b, therefore depend on the nature of the variation in the population. In a genetically homogeneous population h would be zero because genes are unimportant to the observed *variation* within the population. When the genetic deviation and the cultural deviation do not covary ($w = 0$), we have $h^2 + b^2 + e^2 = 1$; in this case, h^2, b^2, and e^2 may be interpreted as the fractions of the variance due to genetic, cultural, and non-transmitted environmental variation, respectively. In studies of quantitative inheritance in which correlations between relatives are analyzed, h^2 is closely related to the *heritability* of the studied character.

The above description depends on the state of the population at a specific point in time. The population changes over time and changes in variance components must be specified in a dynamically consistent manner to obtain a completely rigorous model (Cavalli-Sforza and Feldman, 1973; Feldman and Cavalli-Sforza, 1979; Karlin, 1979). This specification, however, is beyond the scope of the present discussion, and we make the assumption that the population has reached an *equilibrium* so that h, b and

w reflect stable characteristics of the population. As we will see later this assumption entails that w must be specified as a function of h and b.

The model of the individual phenotype given by expression (1) is similar to the description of phenotypic determination in path analysis (see Wright, 1921; Wright, 1931; or Li, 1955). Our model of phenotypic determination corresponds to the path diagram:



and from path analysis we know that the model parameters, the *path coefficients* (h, b and e) are partial regression coefficients. For example, h describes the regression of P on A obtained when the source variables B and E are held constant:

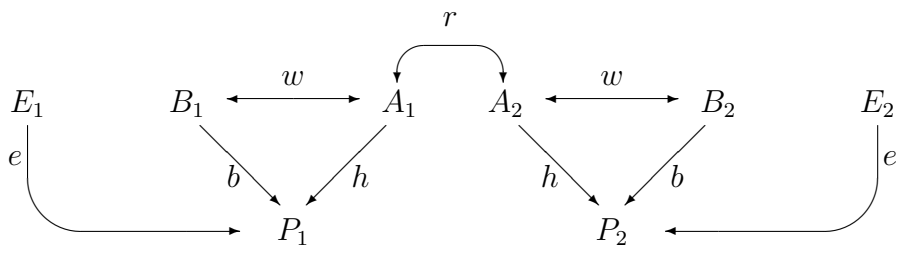
$$h = \frac{\mathrm{Cov}(P, A|B, E)}{\mathrm{Var}(A|B, E)}, \tag{4}$$

and, in a similar way,

$$b = \frac{\mathrm{Cov}(P, B|A, E)}{\mathrm{Var}(B|A, E)} \quad \text{and} \quad e = \mathrm{Cov}(P, E|A, B),$$

because $\mathrm{Var}(E|A, B) = \mathrm{Var}(E) = 1$.

If it is assumed that the only cause of similarity between related individuals arises from their common genes, then the correlation between two related individuals (with phenotypes P_1 and P_2) may be represented by the path diagram



Here we have illustrated the classical assumptions that the only influence that is correlated between the two related individuals is the genetic influence (A_1, A_2) and that the degree of genetic relatedness may be specified by the genetic correlation r (Wright, 1921; Falconer, 1989, p. 166). In this case, the correlation between the two phenotypes is

$$\mathrm{Cov}(P_1, P_2) = \mathrm{Cov}(hA_1 + bB_1 + eE_1, hA_2 + bB_2 + eE_2) = h^2\mathrm{Cov}(A_1, A_2) = h^2 r.$$

Our purpose in this paper is to integrate this approach into a broader framework in which the cultural similarities of related individuals are not ignored.

1.1 Genetic transmission

The genotypic variation in the population is assumed to be stable over time and to originate in allelic variability at a large number of loci not subject to selection. We assume a fully additive genetic model where the genotypic effect is determined by summing the effects of each allele at each locus. The mean genotypic value of an offspring is then simply the average of the parental genotypic values,

$$\mathrm{E}(G_O|G_M, G_F) = \frac{G_M + G_F}{2},$$

where the subscripts, O, M and F, are used to denote the offspring, maternal, and paternal deviations of the relevant variable, here the genotypic value, G. An extension of this analysis would include effects of dominance, epistasis, and genotype-environment interaction (see Falconer, 1989). We will show that the genetic transmission parameter remains $\frac{1}{2}$ even using the normalized genetic measure, A, and even when the parents are related. The assumption of genetic equilibrium entails that the variance in the offspring generation equals the variance in the parental generation, so $\mathrm{Var}(G_O) = \mathrm{Var}(G_M) = \mathrm{Var}(G_F) = V_G$, say. Assuming that the genetic values of the parents are independent, we can write

$$G_O = \frac{G_M + G_F}{2} + D, \tag{5}$$

where D describes the deviation due to Mendelian segregation, i. e. $\mathrm{E}(D) = 0$, $\mathrm{Cov}(D, G_M) = 0$ and $\mathrm{Cov}(D, G_F) = 0$. For the genotypic variance to remain constant

over time with value V_G, equation (5) implies that $\mathrm{Var}(D) = \frac{1}{2}V_G$. Normalizing the variables in (5) by the standard deviation of G, $\sqrt{V_G}$, produces an equation describing the transmission of the normalized genotypic variable $A = G/\sqrt{V_G}$, which was used in the linear model (1):

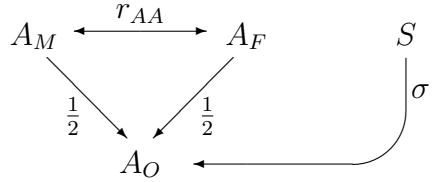
$$A_O = \tfrac{1}{2}A_M + \tfrac{1}{2}A_F + \sigma S, \tag{6}$$

where $S = D/\sqrt{\frac{1}{2}V_G}$ is the normalized segregation variable with $\mathrm{E}(S) = 0$, $\mathrm{Var}(S) = 1$, and $\sigma = \frac{1}{\sqrt{2}}$. The variables A_O, A_M and A_F are normalized variables with mean zero and unit variance, and the segregation variable S is independent of the genotypic values A_M and A_F in the parents.

Now consider the effect of a correlation r_{AA} between the genotypic values of parents. This correlation will occur when the parents are related or when there is assortative mating with respect to the quantitative trait. Assuming that the segregation variance remains the same, $\mathrm{Var}(D) = V_D$, regardless of the value of r_{AA}, a genetic correlation between mates augments the genotypic variance. Using equation (5), we see that

$$\mathrm{Var}(G_O) = \tfrac{1}{4}\mathrm{Var}(G_M) + \tfrac{1}{4}\mathrm{Var}(G_F) + \tfrac{1}{2}r_{AA}\sqrt{\mathrm{Var}(G_M)\mathrm{Var}(G_F)} + V_D,$$

which entails that $V_G = 2V_D/(1-r_{AA})$ at equilibrium. We can use this genetic variance to normalize the genotypic transmission equation (5) and rederive equation (6), but now with $\sigma = \sqrt{\frac{1}{2}(1 - r_{AA})} = \sqrt{1 - \frac{1}{2}(1 + r_{AA})}$. It is therefore reasonable to assume in general that genetic transmission is described by equation (6), or, equivalently, by the path diagram:



(7)

Thus, the assumptions of additive genetics and Mendelian segregation imply that genetic transmission may be described by a single path coefficient equal to $\frac{1}{2}$ independent of the correlation between parents.

1.2 Cultural transmission

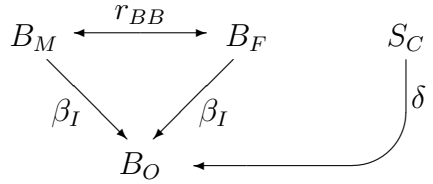
Whereas Mendelian laws of inheritance may be used to justify assumptions about the transmission of genes that affect a continuous trait, no such laws exist for cultural transmission. Rather, the dynamics of cultural transmission will vary depending on the nature of the character under study (Cavalli-Sforza and Feldman, 1981). Here we will limit our attention to vertical transmission, where only parents contribute to the transmissible environmental component of an individual as described by the variable B. Two models that differ in how parents influence the cultural values of their offspring will be examined: an indirect model in which only the cultural values of the parents have an influence ($B \rightarrow B$) and a direct model in which the phenotypes of the parents determine the offspring's cultural value ($P \rightarrow B$).

1.2.1 Indirect cultural transmission

The variable B refers to a latent cultural factor that can be transmitted from generation to generation. This transmission occurs with some error (or cultural "segregation") described by S_C with $\mathrm{E}(S_C) = 0$ and $\mathrm{Var}(S_C) = 1$. That is, we assume a transmission equation for culture similar to equation (6) for genes,

$$B_O = \beta_I(B_M + B_F) + \delta S_C, \tag{8}$$

where β_I describes the fidelity with which family environment is transmitted and δ describes the degree of transmission error. The normalization of the variables produces $\delta = \sqrt{1 - 2\beta_I^2(1 + r_{BB})}$, where r_{BB} is the correlation between the cultural value of mates; $r_{BB} = \mathrm{Cov}(B_M, B_F)$. Thus, indirect cultural transmission may be described by the path diagram:



(9)

When $\beta_I = 1/\sqrt{2(1 + r_{BB})}$, there is no error in the transmission of culture and $\delta = 0$, while if β is less than $1/\sqrt{2(1 + r_{BB})}$ error in transmission occurs, $\delta > 0$, and there

is regression of the cultural value towards the population mean when parents have extreme cultural values.

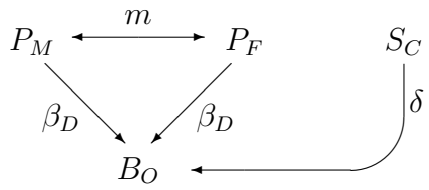
In equation (8), it is assumed that maternal and paternal transmission occur with equal strength. This may be invalid for certain characters (e.g. Cavalli-Sforza and Feldman, 1981; Cavalli-Sforza et al., 1982), but the model may easily be extended to include sex-dependent transmission parameters (see e.g. Karlin, 1979; Feldman and Cavalli-Sforza, 1979; Cloninger et al., 1979a; Rao et al., 1979; Kirkpatrick and Lande, 1989).

1.2.2 Direct cultural transmission

In this case, the variable B describes the cultural experiences of an offspring, but is not itself transmitted. Instead, the cultural value of an offspring B_O is influenced by its parents' phenotypic values and therefore by *all* of the determinants of parental phenotype. The maternal and paternal influences are equal and the cultural transmission occurs with some error, so the cultural value of an offspring is given by:

$$B_O = \beta_D(P_M + P_F) + \delta S_C, \tag{10}$$

where S_C is a normalized variable describing the random error in the cultural transmission. The correlation between the phenotypic values of mates is $m = \mathrm{Cov}(P_M, P_F)$, and since all the variables have unit variance, $\delta = \sqrt{1 - 2\beta_D^2(1+m)}$. When cultural transmission is perfectly faithful, i.e. $\delta = 0$, then β_D equals $1/\sqrt{2(1+m)}$, and if there is error in transmission, i.e. $\delta > 0$, then β_D is less than $1/\sqrt{2(1+m)}$. Equation (10) for direct cultural transmission may be described by the path diagram:



(11)

This model of determination of the cultural value of an individual is considerably simpler than the indirect transmission model given by equation (8) or diagram (9) because the source variables of B_O are observable properties of the parents. This allows cultural transmission to be specified directly in equation (1) for the phenotypic

determination, namely

$$P_O = hA_O + d(P_M + P_F) + e'E'_O. \tag{12}$$

Here the model is specified, without reference to a cultural value, as $d = b\beta_D$ with $d \leq 1/\sqrt{2(1+m)}$, $e'E'_O = eE_O + b\delta S_C$, and $e' = \sqrt{b^2\delta^2 + e^2}$ in terms of the direct transmission model (10). Thus, cultural transmission is fully specified by the coefficient d, so only inferences about the *product* $b\beta_D$ can be made from observations on phenotypic correlations between related individuals. Direct cultural inheritance models may therefore be given in the form (12), without reference to a latent cultural variable, as in the models of Cavalli-Sforza and Feldman (1973), (1978), Feldman and Cavalli-Sforza (1977), (1979) and Martin et al. (1986).

For a full specification of the direct transmission model without reference to a latent cultural value, we will need to specify the correlation structure between the variables used in equation (12). In the next section on assortative mating, we will determine the correlations between A, P, and E in an individual and those values in her or his mate, denoted as r_{AA}, r_{AP} and so on. We will also determine $\gamma = \mathrm{Cov}(P, A)$, which is the correlation between an individual's genetic and phenotypic values, and relate this quantity to w, the correlation between an individual's genetic and cultural values. Substituting the genetic transmission equation (6) into equation (12), γ and the r_{IJ} will be natural quantities in the correlations expected between related individuals. We can also rederive the phenotypic variance in the population without reference to a latent cultural influence, by directly evaluating the variance of equation (12):

$$\mathrm{Var}(P_O) = h^2 + 2d^2(1+m) + (e')^2 + 2hd(\gamma + r_{AP}). \tag{13}$$

It will later be shown that under phenotypic homogamy $\beta_D(\gamma + r_{AP}) = w$. Making this substitution along with the substitutions $d = b\beta_D$, $e' = \sqrt{b^2\delta^2 + e^2}$, and $\delta = \sqrt{1 - 2\beta_D^2(1+m)}$, we see that (13) is equivalent to (2) and hence both can be set to one. We will use equation (13) to find the parameter e' from the remaining parameters. If we also knew the value of e, then we could estimate b separately from β_D. For instance, Eaves et al. (1989) make the assumption that e is zero, allowing them to derive b from the remaining parameters. In our work, we will not make any *a priori*

assumptions concerning the value of e or e'. We will further assume that equation (12) is the more accurate descriptor of phenotypic determination, so that E' and not E will enter into the correlations between relatives. Hence, we will not be able to estimate e, b, or β_D independently, but will be able to estimate d and e'. Provided that the system may be assumed to be at equilibrium, equation (12) is a complete description of phenotypic determination with direct linear inheritance of culture.

1.3 Assortative mating

Empirical studies on human behavioral traits have often revealed strong patterns of non-random mating. One of the most common forms is assortative mating which is a deviation from random mating where similar individuals mate with each other more frequently than would occur at random (Cavalli-Sforza and Bodmer, 1971). Such assortment may be caused by conscious mate choice or by the ways in which humans meet one another. We will examine the effects of two types of assortative mating on linear models of quantitative inheritance: the *phenotypic homogamy* model, where assortment is based on the phenotype of individuals, and the *social homogamy* model, where assortment occurs because individuals belong to the same social group.

Assortment induces correlations among the hidden variables that determine the phenotypes of the mates, and this effect may be described by the variance-covariance matrix of $(A_M, B_M, E_M, A_F, B_F, E_F)$ given by

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & w & 0 & r_{AA} & r_{AB} & r_{AE} \\ w & 1 & 0 & r_{AB} & r_{BB} & r_{BE} \\ 0 & 0 & 1 & r_{AE} & r_{BE} & r_{EE} \\ r_{AA} & r_{AB} & r_{AE} & 1 & w & 0 \\ r_{AB} & r_{BB} & r_{BE} & w & 1 & 0 \\ r_{AE} & r_{BE} & r_{EE} & 0 & 0 & 1 \end{pmatrix}, \tag{14}$$

where r_{IJ} is the correlation between the hidden variable I in the female and the hidden variable J in the male or the other way around. It is convenient to write this matrix

in terms of the variance-covariance matrix $\mathbf{\Omega}$ of (A_M, B_M, E_M) or (A_F, B_F, E_F),

$$\mathbf{\Omega} = \begin{pmatrix} 1 & w & 0 \\ w & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{15}$$

which refers to within-individual variation, and the covariance matrix

$$\mathbf{\Theta} = \begin{pmatrix} r_{AA} & r_{AB} & r_{AE} \\ r_{AB} & r_{BB} & r_{BE} \\ r_{AE} & r_{BE} & r_{EE} \end{pmatrix} \tag{16}$$

that describes the covariances between (A_M, B_M, E_M) and (A_F, B_F, E_F). Equation (14) is then compressed to

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Omega} & \mathbf{\Theta} \\ \mathbf{\Theta} & \mathbf{\Omega} \end{pmatrix}. \tag{17}$$

In the subsequent analysis of assortative mating, we follow closely Goldberger (1978) and Feldman and Cavalli-Sforza (1979) in deriving the correlations between the hidden variables of mates, $\mathbf{\Theta}$, and the correlations between the phenotype of one mate and the latent variables of the other mate:

$$r_{AP} = \mathrm{Cov}(A_M, P_F) = hr_{AA} + br_{AB} + er_{AE} \tag{18}$$

$$r_{BP} = \mathrm{Cov}(B_M, P_F) = hr_{AB} + br_{BB} + er_{BE} \tag{19}$$

$$r_{EP} = \mathrm{Cov}(E_M, P_F) = hr_{AE} + br_{BE} + er_{EE}. \tag{20}$$

from equation (1). The correlations (18) and (20) will be particularly important in the direct transmission model. In this case, the correlations induced by assortative mating are more conveniently described using the variance-covariance matrix of

$(A_M, P_M, E_M, A_F, P_F, E_F)$, given by

$$\tilde{\mathbf{\Gamma}} = \begin{pmatrix} 1 & \gamma & 0 & r_{AA} & r_{AP} & r_{AE} \\ \gamma & 1 & 0 & r_{AP} & m & r_{PE} \\ 0 & 0 & 1 & r_{AE} & r_{PE} & r_{EE} \\ r_{AA} & r_{AP} & r_{AE} & 1 & \gamma & 0 \\ r_{AP} & m & r_{PE} & \gamma & 1 & 0 \\ r_{AE} & r_{PE} & r_{EE} & 0 & 0 & 1 \end{pmatrix}, \tag{21}$$

where γ is the covariance between the genotypic and the phenotypic variables of an individual.

1.3.1 Phenotypic homogamy

The simplest model assumes that assortment is based on phenotype and this assortment creates a correlation $m = \mathrm{Cov}(P_M, P_F)$ between the phenotypic values of mated pairs. That is, phenotypic homogamy occurs when individuals tend to choose mates that have similar trait values.

The model of phenotypic assortment that we consider is equivalent to what Goldberger (1978) calls "Fisher's model" of assortative mating. Here any correlation between the hidden variables A_F, B_F and E_F in the female and hidden variables A_M, B_M and E_M in the male arise only because of their respective correlations with P_F and P_M. That is, we assume that G_M given P_M, written $(G_M|P_M)$, is independent of $(G_F|P_F)$; $(B_M|P_M)$ is independent of $(B_F|P_F)$; and $(E_M|P_M)$ is independent of $(E_F|P_F)$. These assumptions (together with the assumption of a multivariate Gaussian distribution) allow the covariance matrix $\mathbf{\Theta}$ to be written as

$$\mathbf{\Theta} = m \begin{pmatrix} \gamma^2 & \gamma\phi & \gamma e \\ \gamma\phi & \phi^2 & \phi e \\ \gamma e & \phi e & e^2 \end{pmatrix}, \tag{22}$$

where $\gamma = \mathrm{Cov}(A, P)$ and $\phi = \mathrm{Cov}(B, P)$ are covariances between the hidden and the observable variables of an individual. To see that this covariance matrix results from the assumptions, consider, for instance, the covariance between the genotype of

one mate and the cultural value of the other:

$$\begin{aligned}
\mathrm{Cov}(A_M, B_F) &= \mathrm{E}(A_M B_F) \\
&= \mathrm{E}\left[\mathrm{E}(A_M B_F | P_M, P_F)\right] \\
&= \mathrm{E}\left[\mathrm{E}(A_M | P_M)\mathrm{E}(B_F | P_F)\right] \\
&= \mathrm{E}\left[\mathrm{Cov}(A_M, P_M) P_M \,\mathrm{Cov}(B_F, P_F) P_F\right] \\
&= \gamma\phi m,
\end{aligned}$$

where we first used the conditional independence and then a well-known property of the two-dimensional Gaussian distribution. From equation (1) for the phenotype we obtain the within-individual correlations as

$$\gamma = \mathrm{Cov}(A, P) = \mathrm{Cov}(A, hA + bB + eE) = h + bw, \tag{23}$$

$$\phi = \mathrm{Cov}(B, P) = \mathrm{Cov}(B, hA + bB + eE) = hw + b. \tag{24}$$

From these basic relations and $\boldsymbol{\Theta}$ it is possible to derive all of the correlations between the hidden variables (A, B and E) of mates:

$$r_{AA} = \mathrm{Cov}(A_M, A_F) = m\gamma^2 = m(h + bw)^2, \tag{25}$$

$$\begin{aligned}
r_{AB} &= \mathrm{Cov}(A_M, B_F) \\
&= \mathrm{Cov}(B_M, A_F) = m\gamma\phi = m(h + bw)(b + hw),
\end{aligned} \tag{26}$$

$$r_{BB} = \mathrm{Cov}(B_M, B_F) = m\phi^2 = m(b + hw)^2, \tag{27}$$

$$\begin{aligned}
r_{AE} &= \mathrm{Cov}(A_M, E_F) \\
&= \mathrm{Cov}(E_M, A_F) = me\gamma = me(h + bw),
\end{aligned} \tag{28}$$

$$\begin{aligned}
r_{BE} &= \mathrm{Cov}(E_M, B_F) \\
&= \mathrm{Cov}(B_M, E_F) = me\phi = me(b + hw),
\end{aligned} \tag{29}$$

$$r_{EE} = \mathrm{Cov}(E_M, E_F) = me^2. \tag{30}$$

Equation (1) can be used to check that the correlations between the hidden variables [25–30] produce the correlation, m, between the phenotypes of mates. For this we

must show that

$$\begin{aligned} m &= \mathrm{Cov}(P_M, P_F) = \mathrm{Cov}(hA_M + bB_M + eE_M, hA_F + bB_F + eE_F) \\ &= h^2 r_{AA} + b^2 r_{BB} + 2hbr_{AB} + 2her_{AE} + 2ber_{BE} + e^2 r_{EE}, \end{aligned} \tag{31}$$

which follows by substituting e from (3). Thus, with phenotypic assortment, the correlation matrix $\mathbf{\Theta}$ is

$$\mathbf{\Theta} = m \begin{pmatrix} (h+bw)^2 & (h+bw)(b+hw) & (h+bw)e \\ (h+bw)(b+hw) & (b+hw)^2 & (b+hw)e \\ (h+bw)e & (b+hw)e & e^2 \end{pmatrix} \tag{32}$$

in terms of the basic parameters of phenotypic determination.

The correlations (18) - (20) between the phenotype of one mate and the latent variables of the other mate may be obtained from equation (22). Thus,

$$r_{AP} = \mathrm{Cov}(A_M, P_F) = \mathrm{Cov}(A_M, hA_F + bB_F + eE_F) = m\gamma(h\gamma + b\phi + e^2),$$

and using equation (3) for e and the definitions (23) and (24), the last parenthesis must be unity, i.e. $h\gamma + b\phi + e^2 = 1$. Repeating this calculation for r_{BP} and r_{EP} we have

$$r_{AP} = m\gamma, \qquad r_{BP} = m\phi, \quad \text{and} \quad r_{EP} = me. \tag{33}$$

The variance-covariance matrix $\tilde{\mathbf{\Gamma}}$ of $(A_M, P_M, E_M, A_F, P_F, E_F)$ may be constructed by incorporating the above calculations,

$$\tilde{\mathbf{\Gamma}} = \begin{pmatrix} 1 & \gamma & 0 & m\gamma^2 & m\gamma & m\gamma e \\ \gamma & 1 & 0 & m\gamma & m & me \\ 0 & 0 & 1 & m\gamma e & me & me^2 \\ m\gamma^2 & m\gamma & m\gamma e & 1 & \gamma & 0 \\ m\gamma & m & me & \gamma & 1 & 0 \\ m\gamma e & me & me^2 & 0 & 0 & 1 \end{pmatrix}, \tag{34}$$

but it may also be found directly using the assumption that correlations between hidden variables in mates arise only because of their respective correlations with P_F and P_M.

1.3.2 Social homogamy

Non-random mating can also result from assortment on the basis of the cultural value or the genotype of an individual (Feldman and Cavalli-Sforza, 1977, 1979; Rao et al., 1976; Rao et al., 1979; Morton and Rao, 1978; Rao and Morton, 1978). Morton and Rao (1978) described a simple model where the correlation observed between mates arises not through direct phenotypic assortment but through common membership in a social or geographical subset of the population. Assortment may occur on the basis of social class, residence, school, church, job, social activities or race (Morton and Rao, 1978). To the extent that mating occurs among individuals within a group and to the extent that groups differ in their composition with respect to cultural and genetic variables, group membership will produce a correlation between the cultural and genotypic values of mates which will lead, indirectly, to a correlation between the phenotypes of spouses.

Although groups may differ in their genetic and cultural composition, it is assumed that the distribution of non-transmissible environmental experiences is *independent* of the grouping. Specifically, any and all experiences that depend on group membership will be transmissible to the extent that group membership is inherited, and these experiences would be included in the cultural value summarized by the variable B. Therefore, specific environmental experiences summarized in E, which by definition are not transmitted from generation to generation, cannot depend on grouping. Thus, assortment that occurs on the basis of grouping may be specified by the correlation coefficients r_{AA}, r_{AB} and r_{BB} between the variables A and B of mated pairs, while correlations that involve the specific environmental values of mates, i. e. r_{AE}, r_{BE} and r_{EE}, will all be zero:

$$\mathbf{\Theta} = \begin{pmatrix} r_{AA} & r_{AB} & 0 \\ r_{AB} & r_{BB} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{35}$$

This formulation follows that of Goldberger (1978) and Rao et al. (1979).

In this model $m = \mathrm{Cov}(P_M, P_F)$ is a parameter derived from the model parameters r_{AA}, r_{AB} and r_{BB}. Using equation (1) to determine phenotype, the expected correlation between the phenotypic values of mated pairs is

$$m = \mathrm{Cov}\left(hA_M + bB_M + eE_M, hA_F + bB_F + eE_F\right)$$

$$= \quad h^2 r_{AA} + 2hbr_{AB} + b^2 r_{BB}. \tag{36}$$

Thus $h^2 r_{AA} + 2hbr_{AB} + b^2 r_{BB}$ must be less than one. Morton and Rao (1978) and Rao et al. (1979) assume that the relation $r_{AB} = \sqrt{r_{AA} r_{BB}}$ holds, reflecting the assumption that all marital correlations stem from group membership. We may then rewrite equation (36) as

$$m = \left(h\sqrt{r_{AA}} + b\sqrt{r_{BB}}\right)^2.$$

The variance-covariance matrix $\mathbf{\Gamma}$ given in equation (17) is now specified by

$$\mathbf{\Theta} \; = \; m \begin{pmatrix} r_{AA} & \sqrt{r_{AA} r_{BB}} & 0 \\ \sqrt{r_{AA} r_{BB}} & r_{BB} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{37}$$

Repeating the calculations performed to obtain equation (33), the correlations between the phenotype of one mate and the latent variables in the other mate are:

$$r_{AP} = hr_{AA} + br_{AB}, \qquad r_{BP} = hr_{AB} + br_{BB}, \quad \text{and} \quad r_{EP} = 0,$$

for the case of social homogamy.

It is possible, of course that both phenotypic and social assortment occur within a population. Rao et al. (1979), Rao et al. (1982), and Eaves et al. (1989) discuss mixed homogamy models. These models may be used, in theory, to determine the relative importance of the two types of assortative mating, but as we shall see the social homogamy model already specifies enough parameters to use most of the degrees of freedom available in a typical data set.

1.3.3 Regression analysis of in-laws

The linear models described above decompose the phenotype of an individual into components whose transmission from the previous generation is known. However, we have yet to specify how variables in relatives related by marriage (in-laws) can be written in terms of variables in previous generations. In order to determine the correlation between, for instance, individuals and their parents-in-law, we use a method similar to that described in Rao et al. (1979). This method regresses the genetic,

cultural, and environmental values of an individual onto these variables in her or his mate.

With assortative mating, the latent variables of spouses will generally be correlated, with correlation coefficients given by the matrix (32) in the case of phenotypic assortment and by the matrix (37) in the case of social homogamy. Let us write the latent variables of one spouse ($S2$) as linear functions of the latent variables in the other spouse ($S1$):

$$\begin{pmatrix} A_{S2} \\ B_{S2} \\ E_{S2} \end{pmatrix} = \mathbf{\Lambda} \begin{pmatrix} A_{S1} \\ B_{S1} \\ E_{S1} \end{pmatrix} + \begin{pmatrix} \epsilon_1 W_1 \\ \epsilon_2 W_2 \\ \epsilon_3 W_3 \end{pmatrix}, \tag{38}$$

where

$$\mathbf{\Lambda} = \begin{pmatrix} c_{AA} & c_{AB} & c_{AE} \\ c_{BA} & c_{BB} & c_{BE} \\ c_{EA} & c_{EB} & c_{EE} \end{pmatrix} \tag{39}$$

is a matrix of partial regression coefficients and W_i, $i = 1, 2, 3$, are normalized error variables (see Appendix A and especially equation (94)). Specifically, each c_{IJ} measures the degree to which the latent variable I in one spouse depends on the latent variable J and only J in the other spouse. For instance c_{AB} equals:

$$\frac{\text{Cov}(A_F, B_M | A_M, E_M)}{\text{Var}(B_M | A_M, E_M)}.$$

Using equation (38), the matrix of correlations between spouses, $\mathbf{\Theta}$, given earlier as (16), may be obtained in terms of the regression coefficients in Λ as:

$$\mathbf{\Theta} = \mathbf{\Lambda\Omega},$$

where $\mathbf{\Omega}$ is the variance-covariance matrix (15) of the hidden variables in an individual. This equation implies that

$$\mathbf{\Lambda} = \mathbf{\Theta\Omega}^{-1}, \tag{40}$$

and the inverse of the matrix $\mathbf{\Omega}$ is given by

$$\mathbf{\Omega}^{-1} = \frac{1}{1-w^2}\begin{pmatrix} 1 & -w & 0 \\ -w & 1 & 0 \\ 0 & 0 & 1-w^2 \end{pmatrix}.$$

These equations just express the usual relationship between correlation coefficients and partial regression coefficients.

With phenotypic homogamy, where $\mathbf{\Theta}$ is given by (32), we have from equation (40):

$$\mathbf{\Lambda} = m\begin{pmatrix} \gamma h & \gamma b & \gamma e \\ \phi h & \phi b & \phi e \\ he & be & e^2 \end{pmatrix}, \tag{41}$$

and these regression coefficients are consistent with the path diagrams used by Cloninger et al. (1979a). Inserting $\mathbf{\Lambda}$ from equation (41) back into equation (38) yields

$$\begin{pmatrix} A_{S2} \\ B_{S2} \\ E_{S2} \end{pmatrix} = m\begin{pmatrix} \gamma \\ \phi \\ e \end{pmatrix} P_{S1} + \begin{pmatrix} \epsilon_1 W_1 \\ \epsilon_2 W_2 \\ \epsilon_3 W_3 \end{pmatrix} \tag{42}$$

as expected from the conditional independence of the latent variables in the mates. Multiplying equation (42) on the left by the row vector (h, b, e), we obtain

$$P_{S2} = m(h\gamma + b\phi + e^2)P_{S1} + h\epsilon_1 W_1 + b\epsilon_2 W_2 + e\epsilon_3 W_3.$$

Recall that $h\gamma + b\phi + e^2$ equals one and rewrite $h\epsilon_1 W_1 + b\epsilon_2 W_2 + e\epsilon_3 W_3$ as ϵW to obtain:

$$P_{S2} = mP_{S1} + \epsilon W.$$

This formulation is thus consistent with the claim that m may be viewed as a partial regression coefficient in models of phenotypic homogamy (Rao et al., 1979).

With social homogamy, where $\mathbf{\Theta}$ is given by (37), the partial regression coefficients

become:

$$\mathbf{\Lambda} = \frac{1}{1-w^2}\begin{pmatrix} r_{AA}-wr_{AB} & r_{AB}-wr_{AA} & 0 \\ r_{AB}-wr_{BB} & r_{BB}-wr_{AB} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{43}$$

where r_{AB} is assumed to equal $\sqrt{r_{AA}r_{BB}}$ (Rao et al., 1979).

Using equation (38), which specifies the latent variables of an individual as linear functions of the latent variables in her or his mate, and using the above formulae for the partial regression coefficients within these equations, we can now derive the correlations between remote relatives for which relations between in-laws play a role. For instance, the expected correlation between an individual and her or his parent-in-law may be computed and this correlation will in turn be important in deriving the correlation between grandparents and grandoffspring and all other remote relatives.

1.4 Summary of the different approaches

The models we have considered are based on biometric models for the determination of phenotype and genetic transmission, and they may be viewed as simple generalizations of the standard models of quantitative genetics. They are extended by different assumptions about the vertical transmission of cultural attributes of individuals, and incorporate similarities between mates due to assortative mating. The models specify the phenotype of an individual linearly in terms of genotypic and environmental variables, which are in turn specified in terms of the variables in the individual's parents. The main features that distinguish the various models that we analyze are summarized as follows.

The Determination of Phenotype — The determination of phenotype, P, is given by equation (1) as a linear function of latent variables that describe the genotypic, A, the cultural, B, and the specific environmental, E, contributions to the phenotype.

Genetic Transmission — The genotype A is transmitted autonomously between generations in that the offspring genotypic effect is given as a function only of the genotypic effects of the parents, by equation (6). The assumption of multifactorial additive genetic determination entails that the coefficients describing the transmission are independent of the genetic composition of the population and of the mating structure in the population.

Cultural Transmission — Two models are used to describe transmission of the latent cultural value from the parents. With *indirect* transmission, the latent cultural variable B is transmitted in a manner analogous to the transmission of the genotypic effects A, except that B is transmitted from parents to offspring by non-genetic means, equation (8). The observed trait is therefore not the transmitted factor, but some combination of perhaps unobservable factors is transmitted. With *direct* transmission the observed trait values, namely the phenotypes of parents, directly influence the cultural values of their children, according to equation (10). Equivalently, this transmission model may be formulated as a variant of the model for the determination of the phenotype, where the parental phenotypes directly influence the phenotypes of offspring, according to equation (12).

Assortative Mating — *Phenotypic homogamy* is assortment with respect to the character under study, where individuals choose to mate with similar individuals. *Social homogamy* occurs when individuals tend to mate according to the group to which they belong. Because groups may differ in their cultural and genotypic values, group membership may produce correlations between the latent variables (A and B) of mates and this results indirectly in a correlation between their phenotypes (P).

Statistical models — Two models with indirect transmission and one model with direct transmission (Table 1) will be considered in the following analyses of published data. The direct-transmission model will be discussed only in the case of phenotypic homogamy. This model is virtually identical to the model considered by Martin et al. (1986), Cavalli-Sforza and Feldman (1978), and Feldman and Cavalli-Sforza (1979). Both phenotypic and social homogamy will be considered in the case of indirect transmission. The phenotypic-homogamy model was considered by Rice et al. (1978), Cloninger et al. (1979a), Cloninger et al. (1979b) and Feldman and Cavalli-Sforza (1979), while the social-homogamy model is that used by Rao et al. (1976), Morton and Rao (1978), and Rao et al. (1979).

Table 1 here.

These models have been completely specified in the previous sections and we now use them to compute expected correlations between relatives. These expectations will then be compared to observed correlations for various traits, and this comparison will permit estimation of the parameters that define the models.

2 Familial Correlations in Biological Families

Vertical transmission of a trait will produce familial correlations in that trait and the values of these correlations will depend on the details of the transmission. The simplest and most fundamental relation in the study of inheritance is that between a parent and its offspring. In the transmission models described above the parent-offspring relationship is independent of the sex of the parent, so the parent-offspring correlation may be regarded as the covariance between mother and offspring, $r_{OPT} = \mathrm{Cov}(P_M, P_O)$. For both cultural transmission models we may specify the offspring phenotype as

$$P_O = \tfrac{1}{2}h(A_M + A_F) + bB_O + \sigma S + eE_O \tag{44}$$

from equations (1) and (6). Using this specification we then have

$$\begin{aligned} r_{OPT} &= \tfrac{1}{2}h\mathrm{Cov}(P_M, A_M + A_F) + b\mathrm{Cov}(P_M, B_O) \\ &= \tfrac{1}{2}h(\gamma + r_{AP}) + b\mathrm{Cov}(P_M, B_O). \end{aligned} \tag{45}$$

The correlation r_{AP} depends on the mating structure in the population, while the correlation $\mathrm{Cov}(P_M, B_O)$ depends on the mode of cultural transmission.

The parent-offspring correlation simplifies in the classical genetic model with no cultural transmission where $b = 0$. For phenotypic assortment we have $r_{AP} = m\gamma$, and $\gamma = h$ when $b = 0$, so in this simple model we have

$$r_{OPT} = \tfrac{1}{2}h^2(1 + m). \tag{46}$$

This expression is equivalent to the parent-offspring correlation first derived by Fisher (1918). This equivalence is expected because our regression model of genetic transmission (6) is in essence Fisher's model of genetic transmission.

The relationship between sibs is also fundamental in the study of heredity. Sibs share parents, but equally important in the present context they share the childhood environment to an extent that may not be specified by the parental phenotypes or their determinants in equation (1). The expected correlation between two siblings O and O' may be written as

$$\mathrm{Cov}(P_O, P_{O'}) = \mathrm{Cov}(hA_O + bB_O + eE_O, hA_{O'} + bB_{O'} + eE_{O'})$$

$$= \quad \mathrm{Cov}(hA_O + bB_O, hA_{O'} + bB_{O'}) + e^2\mathrm{Cov}(E_O, E_{O'}), \qquad (47)$$

where the environmental correlation, $\mathrm{Cov}(E_O, E_{O'})$, may depend on the nature of the sib relationship: ordinary full sibs, half sibs, dizygous twins or monozygous twins.

With ordinary full siblings, O and O' have the same parents and equation (47) may be rewritten as a variance between variables in the parental generation, using equation (6) for genetic transmission and equation (8) or (10) for cultural transmission. The correlation between full sibs is

$$r_{SST} = \mathrm{Var}\left[\tfrac{1}{2}h(A_M + A_F) + \beta_I b(B_M + B_F)\right] + c_s e^2 \qquad (48)$$

for indirect cultural transmission, and

$$r_{SST} = \mathrm{Var}\left[\tfrac{1}{2}h(A_M + A_F) + \beta_D b(P_M + P_F)\right] + c_s e^2 \qquad (49)$$

for direct transmission, where $c_s = \mathrm{Cov}\left(E_O, E_{O'}\right)$. For the direct transmission model, we have suppressed the use of primes (e', E') introduced in equation (12), but it is assumed that c_s is actually $\mathrm{Cov}\left(E'_O, E'_{O'}\right)$. This implies that either transmission is truly direct from parents' phenotypes or that there is a latent variable B whose transmission error S_C is also correlated among siblings to the same extent, $c_s = \mathrm{Cov}\left(E'_O, E'_{O'}\right) = \mathrm{Cov}\left(E_O, E_{O'}\right) = \mathrm{Cov}\left((S_C)_O, (S_C)_{O'}\right)$. Dizygotic twins have the same correlation as siblings except, possibly, for a different degree of shared environment, so it is reasonable to replace c_s in equations (48) and (49) by the correlation c_{dz}. Monozygous twins are a special category of siblings, since they share the same genes, so $A_O = A_{O'} = A$. As a consequence, the correlation (r_{MZT}) between a pair of monozygous twins becomes

$$r_{MZT} = \mathrm{Cov}(hA + bB_O, hA + bB_{O'}) + c_{mz} e^2. \qquad (50)$$

In the following sections, explicit formulae are provided for parent-offspring and sib-sib correlations for the three models: the two cultural transmission models with phenotypic assortment and the indirect cultural transmission model with social homogamy.

2.1 Indirect cultural transmission

Indirect cultural inheritance as specified in equation (8) is of the same form as the genetic inheritance in equation (6). In the parent-offspring correlation given by (45) the term $b\mathrm{Cov}(P_M, B_O)$ becomes $\beta_I b\mathrm{Cov}(P_M, B_M + B_F)$ and, since $\mathrm{Cov}(P_M, B_M + B_F) = \phi + r_{BP}$,

$$r_{OPT} = \tfrac{1}{2}h(\gamma + r_{AP}) + \beta_I b(\phi + r_{BP}). \tag{51}$$

For the phenotypic assortment model, from equations (33) we have

$$\begin{aligned} r_{OPT} &= (1+m)(\tfrac{1}{2}h\gamma + \beta_I b\phi) \\ &= (1+m)\left[\tfrac{1}{2}h(h+bw) + \beta_I b(hw+b)\right], \end{aligned} \tag{52}$$

using γ and ϕ from equations (23) and (24). For the social homogamy model, from (51) with (35) we have

$$r_{OPT} = \tfrac{1}{2}h(h + bw + hr_{AA} + br_{AB}) + \beta_I b(hw + b + hr_{AB} + br_{BB}). \tag{53}$$

The full-sib correlation is given by equation (48) which expands as

$$r_{SST} = 2(\tfrac{1}{4}h^2 + \beta_I^2 b^2) + 2h\beta_I bw + 2(r_{AA}\tfrac{1}{4}h^2 + r_{BB}\beta_I^2 b^2) + 2r_{AB}h\beta_I b + c_s e^2.$$

This form of the correlation is already written in terms of the variables of the social homogamy model, and upon rearrangement we have

$$r_{SST} = \tfrac{1}{2}h^2(1 + r_{AA}) + 2\beta_I^2 b^2(1 + r_{BB}) + 2h\beta_I b(w + r_{AB}) + c_s e^2. \tag{54}$$

With phenotypic assortment, however, a further specification is possible by introducing the marital correlations r_{AA}, r_{AB}, and r_{BB} from equation (22):

$$r_{SST} = \tfrac{1}{2}h^2(1 + m\gamma^2) + 2\beta_I^2 b^2(1 + m\phi^2) + 2h\beta_I b(w + m\gamma\phi) + c_s e^2. \tag{55}$$

Similarly, the correlations for monozygous twins (50) for social homogamy and phenotypic homogamy, respectively, are

$$r_{MZT} = h^2 + 2hbw + 2(1 + r_{BB})\beta_I^2 b^2 + c_{mz} e^2 \tag{56}$$

and

$$r_{MZT} = h^2 + 2hbw + 2(1 + m\phi^2)\beta_I^2 b^2 + c_{mz}e^2. \tag{57}$$

2.1.1 Equilibrium gene-culture correlation with indirect transmission

The correlation between the genotype and the cultural value of an individual, w, is specified by the assumptions of the model. A recurrence equation for w may be obtained by considering the correlation w' between A_O and B_O:

$$\begin{aligned} w' &= \mathrm{Cov}\left[\tfrac{1}{2}(A_M + A_F), \beta_I(B_M + B_F)\right] \\ &= \beta_I(w + r_{AB}). \end{aligned} \tag{58}$$

In the absence of assortative mating we have $r_{AB} = 0$ and w will decrease towards zero at the rate β_I. On the other hand when $r_{AB} \neq 0$ assortative mating results in an equilibrium with $w \neq 0$, and a correlation between the genotype and the cultural environment of an individual is maintained.

In the phenotypic assortment model r_{AB} depends on w, so equation (58) may be made explicit by using equation (26) for r_{AB}:

$$w' = \beta_I \left[w + m(h + bw)(b + hw)\right]. \tag{59}$$

Thus, w converges to a unique equilibrium value that satisfies the quadratic equation:

$$w^2 hbm\beta_I + w(\beta_I + \beta_I m b^2 + \beta_I m h^2 - 1) + hbm\beta_I = 0 \tag{60}$$

when

$$\beta_I \left(1 + m(h + b)^2\right) < 1, \tag{61}$$

and otherwise the model degenerates.

The parameters of the phenotypic assortment model are the coefficients h and b describing phenotypic determination (1), the coefficient β_I of the transmission model (9), the correlation m between mated pairs, and the correlations c_s, c_{dz}, and c_{mz} between the non-transmitted environments of siblings. Later, in the analysis of data, the population will be assumed to be at equilibrium, i. e. the value of w satisfies (60).

With social homogamy equation (58) is considerably simpler in that r_{AB} is independent of w and equal to $\sqrt{r_{AA}r_{BB}}$:

$$w' = \beta_I(w + \sqrt{r_{AA}r_{BB}}).$$

Thus, w converges at a rate β_I to

$$w = \frac{\beta_I\sqrt{r_{AA}r_{BB}}}{1-\beta_I}. \tag{62}$$

The parameters for the social homogamy model are h, b, β_I, r_{AA}, r_{BB}, c_s, c_{dz}, and c_{mz}, and as before we will assume that the population has reached the equilibrium (62).

2.2 Direct cultural transmission

The direct cultural transmission model (10), where the cultural variable, B, is specified directly by the phenotypes of the parents, will be considered only for phenotypic assortative mating (Table 1). Then we need not postulate the existence of a latent cultural variable either for transmission or for assortative mating. Under this model, the parent-offspring correlation (45) is given by:

$$\begin{aligned} r_{OPT} &= \tfrac{1}{2}h(\gamma + r_{AP}) + \beta_D b(1+m) \\ &= (1+m)(\tfrac{1}{2}h\gamma + \beta_D b). \end{aligned} \tag{63}$$

The correlation between two siblings is given by equation (49) which may be expanded to become

$$r_{SST} = \tfrac{1}{2}(1+r_{AA})h^2 + 2(\beta_D b)^2(1+m) + 2\beta_D bh(\gamma + r_{AP}) + c_s e^2$$

which, using (34), becomes

$$r_{SST} = \tfrac{1}{2}(1+m\gamma^2)h^2 + 2(\beta_D b)^2(1+m) + 2\beta_D bh\gamma(1+m) + c_s e^2, \tag{64}$$

while the correlation between a pair of monozygous twins given by equation (50) becomes

$$r_{MZT} = h^2 + 2whb + 2(1+m)(\beta_D b)^2 + c_{mz}e^2. \tag{65}$$

The correlations between siblings (r_{SST} and r_{MZT}) at equilibrium are equivalent to those given in Martin et al. (1986). They use μ for m, b for $\beta_D b$, and zero for c_s, c_{dz}, and c_{mz}.

2.2.1 Equilibrium genotype-phenotype correlation with direct transmission

For a full specification of the model, the correlation between the variables that determine phenotype, (1) or (12), must also be described. It is convenient in the direct inheritance model to follow the correlation between the genotype and phenotype of an individual, γ. The correlation γ' between the genotypic A_O and phenotypic P_O values in the offspring generation is:

$$\begin{aligned} \gamma' &= \mathrm{Cov}(A_O, hA_O + bB_O) \\ &= h + b\mathrm{Cov}\left[\tfrac{1}{2}(A_M + A_F), \beta_D(P_M + P_F)\right]. \end{aligned}$$

Under phenotypic homogamy, using equation (34), this gives a recurrence equation in γ,

$$\gamma' = h + \beta_D b\gamma(1 + m). \tag{66}$$

The equilibrium genotype-phenotype correlation is then

$$\gamma = \frac{h}{1 - \beta_D b(1 + m)}, \tag{67}$$

to which convergence occurs at the rate $\beta_D b(1 + m)$.

At equilibrium the correlation between the genotypic A_O and cultural B_O values of an individual is given by

$$\begin{aligned} w &= \mathrm{Cov}(A_O, B_O) \\ &= \mathrm{Cov}\left[\tfrac{1}{2}(A_M + A_F), \beta_D(P_M + P_F)\right] \\ &= \beta_D\gamma(1 + m). \end{aligned} \tag{68}$$

w will only enter the familial correlations as a part of the product bw, which can be rewritten as $d\gamma(1 + m)$ where $d = \beta_D b$. This emphasizes that only d is estimable, since b and β_D do not enter the equations independently.

The parameters of the direct cultural transmission model with phenotypic assortment are the phenotypic determination coefficient h, the product d between the phenotypic determination coefficient b and the transmission coefficient β_D, the mating correlation m, and the environmental correlations c_s, c_{dz}, and c_{mz}.

2.3 Remote relatives

As an example of the calculation of correlations between the phenotypes of remote relatives, consider the correlation (r_{GPO}) between a maternal grandfather (GF) and a grandoffspring (O). In this case, the father is the son-in-law of the grandfather and so the regression equations given in section 1.3.3 must be used. The grandparent-grandoffspring correlation may then be calculated in the indirect inheritance model as follows:

$$\begin{aligned} r_{GPO} &= \mathrm{Cov}(P_{GF}, P_O) \\ &= \mathrm{Cov}\left[P_{GF}, \tfrac{1}{2}h(A_M + A_F) + b\beta_I(B_M + B_F)\right] \\ &= \mathrm{Cov}\Big[P_{GF}, \tfrac{1}{2}h(A_M + c_{AA}A_M + c_{AB}B_M + c_{AE}E_M) \\ &\qquad + b\beta_I(B_M + c_{BA}A_M + c_{BB}B_M + c_{BE}E_M)\Big]. \end{aligned}$$

This expression involves only the attributes of the maternal grandfather and the mother, i.e. attributes of a parent and an offspring. We may therefore rewrite this as:

$$\begin{aligned} r_{GPO} &= \mathrm{Cov}\Big[P_M, \tfrac{1}{2}h(A_O + c_{AA}A_O + c_{AB}B_O + c_{AE}E_O) \\ &\qquad + b\beta_I(B_O + c_{BA}A_O + c_{BB}B_O + c_{BE}E_O)\Big] \\ &= \mathrm{Cov}\Big[hA_M + bB_M + eE_M, \tfrac{1}{4}h(1 + c_{AA})(A_F + A_M) \\ &\qquad + \tfrac{1}{2}hc_{AB}\beta_I(B_F + B_M) + b\beta_I^2(1 + c_{BB})(B_F + B_M) \\ &\qquad + \tfrac{1}{2}b\beta_I c_{BA}(A_F + A_M)\Big] \\ &= [\tfrac{1}{4}h^2(1 + c_{AA}) + \tfrac{1}{2}hb\beta_I c_{BA}](r_{AA} + 1) \\ &\quad + [\tfrac{1}{2}h^2 c_{AB}\beta_I + hb\beta_I^2(1 + c_{BB})](r_{AB} + w) \\ &\quad + [\tfrac{1}{4}hb(1 + c_{AA}) + \tfrac{1}{2}b^2\beta_I c_{BA}](r_{AB} + w) \\ &\quad + [\tfrac{1}{2}hbc_{AB}\beta_I + b^2\beta_I^2(1 + c_{BB})](r_{BB} + 1) \\ &\quad + [\tfrac{1}{4}eh(1 + c_{AA}) + \tfrac{1}{2}be\beta_I c_{BA}]r_{AE} \end{aligned}$$

$$+[\tfrac{1}{2}hec_{AB}\beta_I + be\beta_I^2(1+c_{BB})]r_{BE}.$$

After some algebraic rearrangement, this reduces to

$$\begin{aligned} r_{GPO} &= [\tfrac{1}{2}h(1+c_{AA}) + b\beta_I c_{BA}]\tfrac{1}{2}[h(r_{AA}+1) + b(r_{AB}+w) + er_{AE}] \\ &\quad +[b\beta_I(1+c_{BB}) + \tfrac{1}{2}hc_{AB}]\beta_I[h(r_{AB}+w) + b(r_{BB}+1) + er_{BE}]. \quad (69) \end{aligned}$$

For phenotypic homogamy the correlations (r_{IJ}'s) are inserted from equation (22) or equation (32) and the regression coefficients (c_{IJ}'s) from equation (41). For social homogamy correlations and regression coefficients are obtained from equations (37) and (43).

For the direct inheritance model the correlation between an individual and her or his maternal grandfather is:

$$r_{GPO} = \mathrm{Cov}\left[P_{GF}, \tfrac{1}{2}h(A_M + A_F) + b\beta_D(P_M + P_F)\right].$$

The variable A_F may be specified from the regression equation (38), and as shown before P_F may be given as mP_M plus an error term. Dropping one generation as before and noting that there will be no covariance with the error terms, this substitution produces:

$$\begin{aligned} r_{GPO} &= \mathrm{Cov}[P_M, \tfrac{1}{2}h(A_O + c_{AA}A_O + c_{AB}B_O + c_{AE}E_O) + b\beta_D(1+m)P_O] \\ &= \mathrm{Cov}[P_M, \tfrac{1}{2}h(A_O + m\gamma hA_O + m\gamma bB_O + m\gamma eE_O) + b\beta_D(1+m)P_O], \end{aligned}$$

where we inserted the regression coefficients from equation (41). From equation (1) we have $hA_O + bB_O + eE_O = P_O$ and

$$\begin{aligned} r_{GPO} &= \mathrm{Cov}[P_M, \tfrac{1}{2}h(A_O + m\gamma P_O) + b\beta_D(1+m)P_O] \\ &= \tfrac{1}{2}h\mathrm{Cov}(P_M, A_O) + [\tfrac{1}{2}hm\gamma + b\beta_D(1+m)]\mathrm{Cov}(P_M, P_O). \end{aligned}$$

Referring to equation (63) this reduces to

$$r_{GPO} = (1+m)\Big\{\tfrac{1}{4}h\gamma + [\tfrac{1}{2}h\gamma m + b\beta_D(1+m)](\tfrac{1}{2}h\gamma + b\beta_D)\Big\}. \quad (70)$$

The method used above to calculate the correlations between the phenotypes of

close relatives may be extended to calculate the expected correlations between any other family members. To determine the covariance between relatives, each latent variable is specified by the variables in the preceding generation in an iterative manner until the common ancestor(s) is reached. This process may be visualized using path diagrams as in Appendix A.

3 Familial Correlations and Adoptions

Vertical cultural transmission produces familial correlations between individuals that are not biologically related, and this phenomenon is especially relevant for families with adopted children. The determination of the phenotype of a child adopted at birth differs from that of a child raised with its biological parents. For an adopted child, the genotypic value A is determined exclusively by its biological parents, whereas the cultural value B is supposed to be determined exclusively by its foster or adoptive parents. The two sets of parents are assumed to be independent, so that the genotypic value and the cultural value of an adoptee are independent. In modern practice, considerable matching of the child and its adoptive parents takes place and this non-randomness necessitates complicated modifications to the following analysis. The basic dependence of the phenotype on the genotypic, cultural and environmental values is no longer given by (1), since A, B, and E are all uncorrelated in an adopted offspring. Therefore, we have

$$\mathrm{Var}(hA^a + bB^a + eE^a) = h^2 + b^2 + e^2 = 1 - 2whb, \tag{71}$$

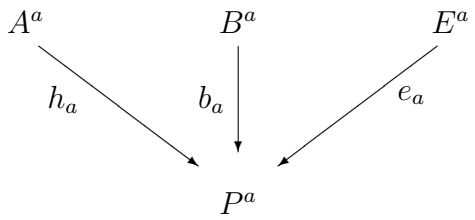
from equation (3), where the superscript "a" emphasizes that the latent variables pertains to an individual adopted away at birth. The phenotypic variance among adoptees is therefore smaller than among non-adopted individuals. Thus, after normalization the equation determining the phenotype of an adopted individual becomes

$$P^a = h_a A^a + b_a B^a + e_a E^a, \tag{72}$$

where $\mathrm{Var}(P^a) = 1$,

$$h_a = \frac{h}{\sqrt{1-2whb}}, \quad b_a = \frac{b}{\sqrt{1-2whb}} \quad \text{and} \quad e_a = \frac{e}{\sqrt{1-2whb}}. \tag{73}$$

The path diagram for the determination of the phenotype of an adoptee therefore includes three independent causes of variation in the relative phenotypic deviation:



Two kinds of familial correlations are considered for adoptees, those involving their biological relatives and those involving the foster family. For parent-offspring correlations these are respectively

$$r_{OPA} = \mathrm{Cov}(P_M^b, P_O^a) = h_a \mathrm{Cov}(P_M, A_O), \tag{74}$$

$$r_{FOP} = \mathrm{Cov}(P_M^f, P_O^a) = b_a \mathrm{Cov}(P_M, B_O), \tag{75}$$

where the superscript "b" refers to a biological parent or relative, the superscript "f" refers to a foster or adoptive parent, and the superscript "a" refers to an adopted offspring. These correlations correspond to each of the two terms in equation (45), and the parent-offspring correlations in the indirect inheritance model reduce to

$$r_{OPA} = \frac{(1+m)\frac{1}{2}h\gamma}{\sqrt{1-2whb}} \tag{76}$$

$$r_{FOP} = \frac{(1+m)\beta_I b\phi}{\sqrt{1-2whb}}, \tag{77}$$

for phenotypic homogamy (see equation 52) and to

$$r_{OPA} = \frac{\frac{1}{2}h\left[h(1+r_{AA})+b(w+r_{AB})\right]}{\sqrt{1-2whb}} \tag{78}$$

$$r_{FOP} = \frac{\beta_I b\left[b(1+r_{BB})+h(w+r_{AB})\right]}{\sqrt{1-2whb}}, \tag{79}$$

for social homogamy (see equation 53). For the direct inheritance model with phenotypic assortment these become

$$r_{OPA} = \frac{(1+m)\frac{1}{2}h\gamma}{\sqrt{1-2(1+m)h\beta_D b\gamma}} \tag{80}$$

$$r_{FOP} = \frac{(1+m)\beta_D b}{\sqrt{1-2(1+m)h\beta_D b\gamma}}, \tag{81}$$

(see equation 63), where the expression (68) has been inserted for w in the denominator.

The correlations between sibs may be of several different types of which we list three that are often used: the correlation between two full sibs who are both adopted away:

$$r_{SSA} = \mathrm{Cov}(P^a_O, P^a_{O'}) = h^2_a \mathrm{Cov}(A_O, A_{O'}),$$

the correlation between an individual in its biological family and a foster sib:

$$r_{SFB} = \mathrm{Cov}(P^b_O, P^f_{O'}) = hb_a \mathrm{Cov}(A_O, B_{O'}) + bb_a \mathrm{Cov}(B_O, B_{O'}) + ee_a \mathrm{Cov}(E_O, E_{O'}),$$

and the correlation between unrelated foster sibs in the same home:

$$r_{SFF} = \mathrm{Cov}(P^f_O, P^f_{O'}) = b^2_a \mathrm{Cov}(B_O, B_{O'}) + e^2_a \mathrm{Cov}(E_O, E_{O'}).$$

These correlations follow from the correlations in Sections 2.1 and 2.2 in the same way as the parent-offspring correlations. These and most of the commonly used correlations are given for the three transmission and mating models in Appendix B.

4 Models and Numbers

Data on the familial aggregation of human behavioral traits have usually been collected in order to elucidate the roles of genetic and cultural transmission in the etiology of the character under study. Biometrical analyses based on the principles outlined in this paper have been used to justify statements about the importance of genetic or cultural transmission for the variation of the character observed in the population. These statements are often framed in terms of a statistic called heritability

which measures the fraction of the total variation that can be said to be genetic in origin.

4.1 The measurement of heritability

The common practical definition of genetic heritability is the variance in the breeding values of individuals, expressed as a fraction of the total phenotypic variance (see Falconer, 1989). This definition immediately links the heritability of a character to the parent-offspring correlation, because by definition the breeding value of an individual is an assigned value that predicts the phenotypic value of the offspring of that individual. Therefore the heritability is often defined as $2b_{OP}$, where b_{OP} is the slope of the parent-offspring regression (see Falconer, 1989).

A more technical genetic definition of heritability is the fraction of the total variance that is attributable to variation in so-called additive effects of alleles (see Ewens, 1979). This definition is rather convenient in theoretical work, but it refers to allelic variability which is unobservable in practice. The beauty of quantitative genetics is that the two definitions of heritability are equivalent in a random mating population. They differ, however, in populations with assortative mating or inbreeding.

The model of multifactorial inheritance that we have used, due to R. A. Fisher (1918), is based on an assumption of additive action of independent single loci. That is, the properties of the quantitative genetic model are simple generalizations based on models of the variation at a single locus. In a one-locus model of assortative mating the parent-offspring covariance takes the value

$$\mathrm{Cov}_{OP} = \tfrac{1}{2}(1+F)V_A, \tag{82}$$

where V_A is the variance due to additive genetic effects and F is Wright's measure of the relative deviation from the Hardy-Weinberg proportions due to assortative mating. The additive variance V_A is itself a function of F; in fact, the additive genetic variance increases with the rate of assortative mating (Fisher, 1918), but that is immaterial for the present argument. Equation (82) shows that $2\mathrm{Cov}_{OP}$ and V_A cannot be the same fractions of the total variance unless $F = 0$. Thus, with assortative mating the practical heritability is different from the theoretically convenient heritability. Comparison of equations (82) and (46) suggests that the quantity h^2 is closely related

to the heritability defined in terms of the additive effect of alleles.

In the classical genetic model with $b = 0$ the full-sib correlation is obtained from equation (55) or (64) as

$$r_{SST} = \tfrac{1}{2}h^2(1 + mh^2). \tag{83}$$

For random mating, $r_{SST} = r_{OPT} = \frac{1}{2}h^2$ in our fully additive model, but with assortative mating the two correlations differ. The difference is that the parent-offspring correlation is modified by the phenotypic correlation between mates through the factor $1+m$, whereas the sib correlation is modified by the genotypic correlation between mates through the factor $1+mh^2$. Thus, the various familial correlations are modified in different ways by assortative mating (see Falconer, 1989 or Ewens, 1979).

Both the practical and technical definitions of heritability that have been discussed so far are measures of the fraction of the phenotypic variance due to the additive effects of alleles, known as the *narrow-sense heritability*. The *broad-sense heritability* is the genotypic variance as a fraction of the phenotypic variance, including non-additive genetic effects. In the strictly additive models that we have considered the variance due to allelic effects is presumed to equal the genotypic variance. This relegates any effects of dominance and epistasis to other variance components; we will return to a discussion of this source of bias in the analysis of data.

The ambiguity in the definition of heritability is not reduced in the presence of cultural transmission. It is natural in our models to treat h^2 as a measure of heritability, but there is no simple relationship between h^2 and the narrow-sense heritability, when h is estimated from several familial correlations in the presence of cultural transmission. Falconer (1989) showed that the narrow-sense heritability increases with the rate of assortative mating. He noted, however, that assortative mating would have a different effect on heritability "thought of as the determination of the resemblance between relatives" (h^2, in our models). We have found in all data sets tested that whenever h is appreciable the estimate for h^2 decreases with increasing rates of assortative mating for the phenotypic homogamy models with indirect (IP) and direct (DP) cultural inheritance (the indirect inheritance model with social homogamy (IS) does not behave in a consistent manner with increasing m). That is, for a particular data set, with higher values of m, a better fit between the model and the data is obtained with lower values of h as expected from equations (46) and (83).

4.2 Heritability measures

The classical biometric model within our modeling framework may be formulated as

$$P = hA + eE$$

with the assumption of a fully additive genetic model and with transmission given by equation (6). In this simple model the breeding value of an individual is given by hA, so that the variance in breeding values (and the heritability) is given by h^2. We may obtain the same estimate for heritability by doubling the parent-offspring regression coefficient ($2b_{OP}$) from the regression equation with random mating

$$\mathrm{E}(P_O|P_M) = b_{OP}P_M = \tfrac{1}{2}h^2P_M. \tag{84}$$

The parent-offspring regression (84) may be used to estimate heritability provided that genetic similarity is the only source of covariation between parents and offspring. This is a classical problem in quantitative genetics, and one way around it has been to standardize the environment to control for other sources of parent-offspring similarity. While culture cannot be standardized with human data, we can calculate the expected heritability if we were able to keep cultural influences constant by considering the conditional parent-offspring regression equation given the parents' and offsprings' cultural values:

$$\mathrm{E}(P_O|P_M, B_O, B_M) = b_{OP|\mathrm{culture}}P_M = \frac{\mathrm{Cov}(P_O, P_M|B_O, B_M)}{\mathrm{Var}(P_M|B_O, B_M)}P_M. \tag{85}$$

Thus, an estimate for the genetic heritability would be twice the quantity

$$b_{OP|\mathrm{culture}} = \frac{\mathrm{Cov}(P_O, P_M|B_O, B_M)}{\mathrm{Var}(P_M|B_O, B_M)} \tag{86}$$

by analogy to equation (84). This equation estimates the contribution of genes to the phenotypic variance if there were no cultural variation, presuming, of course, that the model were correct. It will thus overestimate genetic heritability when cultural influences do contribute to phenotypic variation. In a similar way we may define a

measure of cultural heritability as β^{-1} times the conditional regression coefficient

$$b_{OP|\text{genetics}} = \frac{\text{Cov}(P_O, P_M|A_O, A_M)}{\text{Var}(P_M|A_O, A_M)}, \tag{87}$$

where we now hold constant the genetic effects and examine the influence of culture. Therefore, motivated by these conditional regressions, an estimate of genetic heritability is $2 \times b_{OP|\text{genetics}}$ and that of cultural heritability is ${\beta_I}^{-1} \times b_{OP|\text{genetics}}$.

Our assumption of a multivariate Gaussian distribution of the descriptive variables ensures that the conditional regression coefficients are independent of the particular values given for the cultural variables (in equation 86) or the genotypic variables (in equation 87). Therefore, the conditional heritabilities are in this sense well defined. Since we only consider vertical transmission of culture, the conditional regression coefficients may be viewed as conditional on the parental variables, i. e.

$$\text{Cov}(P_M, P_O|B_O, B_M) = \text{Cov}(P_M, P_O|B_M, B_F).$$

For the indirect inheritance model, (8), with phenotypic assortment, we show in Appendix (C) that:

$$b_{OP|\text{culture}} = \frac{\frac{1}{2}h^2\left(1 + m\gamma^2 - \frac{(w+m\gamma\phi)^2}{1+m\phi^2}\right) + \frac{1}{2}h^2e^2m\left(1 - \frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma\phi^3w}{1-m^2\phi^4}\right)}{h^2\left(1 - \frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma\phi^3w}{1-m^2\phi^4}\right) + e^2}, \tag{88}$$

$$b_{OP|\text{genetics}} = \frac{\beta_I b^2\left(1 + m\phi^2 - \frac{(w+m\gamma\phi)^2}{1+m\gamma^2}\right) + \beta_I b^2e^2m\left(1 - \frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma^3\phi w}{1-m^2\gamma^4}\right)}{b^2\left(1 - \frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma^3\phi w}{1-m^2\gamma^4}\right) + e^2}. \tag{89}$$

This definition of genetic heritability coincides with the classical definition when $b \to 0$, in that $b_{OP|\text{culture}}$ tends to $\frac{1}{2}h^2(1 + mh^2 + me^2) = \frac{1}{2}h^2(1 + m)$ the value in equation (46). Here we have used the equilibrium value of w that approaches zero as $b \to 0$.

A fair approximation of equations (88) and (89) when $m\gamma\phi$ is not too large (hence,

from (59), w is not too large) is

$$\begin{aligned} b_{OP|\text{culture}} &\approx \tfrac{1}{2}h^2\Big[1+m(\gamma^2+e^2)\Big], \\ b_{OP|\text{genetics}} &\approx \beta_I b^2\Big[1+m(\phi^2+e^2)\Big] \end{aligned}$$

where terms of the order $(m\gamma\phi)^2$ have been neglected . If assortative mating is positive, $2b_{OP|\text{culture}} > h^2$. Hence, if we were able to remove the contribution of cultural differences to phenotypic variance, heritability would rise. This illustrates the critical point that any estimate of heritability is sensitive to the cultural context in which the data are obtained.

Similar conditional heritabilities may be calculated for the social homogamy model, but general expressions for the values will not be presented here. For the direct inheritance model, however, conditional heritabilities seem not to be natural quantities, because the phenotypic value of a parent and not just its cultural value constitutes an integral part of the cultural environment of the child.

4.3 Data analysis

The three models in Table 1 will be used to help interpret a variety of data on behavioral characters from human populations. Our analyses will be limited to published estimates of familial correlations (r) and the estimation procedure we use will rely upon Fisher's z-transform of the correlations determined from n pairs of relatives:

$$z = \tfrac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \sim N\left(\zeta, \frac{1}{n-3}\right), \tag{90}$$

where z approximates a normal distribution with a mean of ζ and a standard error of $1/\sqrt{n-3}$ (Fisher, 1993, p.199). We use a least squares method of estimation, in which the parameters are varied until the expected correlations $r^{(i)}_{\text{exp}}$ from the model show minimal deviation from the observed correlations $r^{(i)}_{\text{obs}}$ in the sense that the quantity

$$\text{SSD} = \sum_{i=1}^{N}(n_i-3)\left(z^{(i)}_{obs}-z^{(i)}_{exp}\right)^2 \tag{91}$$

is minimized. Here N is the number of observed correlations and n_i is the number of pairs used in calculating each correlation. This method has been used extensively, e. g. by Rao et al. (1976) and Cloninger et al. (1979a).

The distribution of SSD is expected to approximate a χ^2 distribution with $N - p$ degrees of freedom, where p is the number of estimated parameters (Sokal and Rohlf, 1981). This approximation may be suspect since the z-transforms for different relatives will not be independent (Karlin et al., 1983; Eaves et al., 1989). Thus the results of the following analyses should be interpreted with caution. However, these statistical problems may not be the main caveats to the procedure, because the underlying theoretical models producing the expected values are extremely simplified. Nevertheless, this procedure provides a simple and easy yardstick, and as such has been the statistical procedure most commonly employed in discussions of data on familial correlations in human populations (see also Goldberger, 1978).

The estimation procedure attempts to fit a multi-parameter model to a set of data for which the model provides a rough description. A search procedure similar to "simulated annealing" was employed to find the parameter values giving the smallest SSD. The smoothness of the surface defined by the SSD in equation (91) was checked by using different initial parameter values. In general a unique minimum was found, and Table 2 shows the parameter estimates obtained from three different starting positions when the indirect inheritance, phenotypic homogamy model (the IP Model) was fitted to the data reviewed by Bouchard and McGue (1981), which we discuss in the following section. Table 2 shows that the program gives a reliable fit to the data in that the estimates vary by less than 0.02. Higher reliability can be obtained with longer computation time. Our procedure was able to reproduce independently the parameter estimates and the SSD quoted by Cloninger et al. (1979a) on a subset of the data compiled by Bouchard and McGue (1981).

Table 2 here.

4.4 Analysis of data on IQ variation

Bouchard and McGue (1981) summarized several published studies that reported familial correlations for the intelligence quotient, IQ. Their mean values for the correlations between relatives are given in Table 3. These mean correlations are weighted

averages of correlations from 111 studies on IQ variation. Table 4 records the parameter values which best fit these observed correlations for the three models of Table 1. The heritability estimate ($\hat{h}^2$, last row of table) is highest for the IS Model (indirect transmission, social homogamy), followed by the IP Model (indirect transmission, phenotypic homogamy), with the lowest estimate for the DP Model (direct transmission, phenotypic homogamy). As the χ^2 values show, however, none of the models adequately fits the data.

Tables 3 and 4 here.

The variation of SSD as a function of h may be illustrated by the SSD profile shown in Figure 1 for the IP Model. The profile is constructed by varying all parameters except h so as to minimize the SSD, and the values in the figure are based on the lowest SSD for each of nine values of h (0.1 to 0.9 in increments of 0.1). The SSD profile rises rapidly away from the minimum, which indicates that a good fit is obtained only within a small range of h. The SSD profile may be used to test the null hypothesis that the full model with all parameters fits better than the model with h set to a particular value. The test statistic is the difference between the SSD values for the two models, and is compared to the χ^2_1 distribution (Rao et al., 1974). The interval of h values whose profile SSD is within $\chi^2_1(0.05)$ of the minimum SSD therefore forms an approximate 95% confidence interval for h. In the figure, the dashed line delimits this bound for $\chi^2_1(0.05)$, and only h values between 0.53 and 0.62 have a sufficiently low SSD value to be included in the confidence interval. The confidence limits for h^2 obtained by this method are:

$$\begin{aligned}
&\text{IP Model:} \quad 0.28 < h^2 < 0.38 \quad (\hat{h}^2 = 0.33),\\
&\text{DP Model:} \quad 0.25 < h^2 < 0.35 \quad (\hat{h}^2 = 0.29),\\
&\text{IS Model:} \quad 0.35 < h^2 < 0.48 \quad (\hat{h}^2 = 0.42).
\end{aligned}$$

The two phenotypic homogamy models give very similar estimates of the heritability, and these are somewhat lower than the estimate in the IS Model. The DP Model and the IS Model even give non-overlapping estimates for heritability. These confidence limits, however, should not be too strictly interpreted. They are approximations,

in that the SSD does not actually have a χ^2 distribution, and there is no reason to expect that the correlations are all independent. The relation of the above confidence limits to the genetic influence on the observed variation in IQ depends very much on the ability of the transmission models in section 1 to describe the determination of the IQ of an individual. Thus, we need an evaluation of our confidence in the models themselves.

One assumption made in every model is that the data are homogeneous, that is every subset of the data is described by the same model. In other words, the amount and quality of variation due to genetic causes is assumed to be the same for all classes of individuals. The individuals studied, however, come from a heterogeneous set of environments. It may be, for instance, that twins or relatives raised apart are unusual probands. Removing these subsets from the data we obtain the estimates for h^2 shown in Table 5. In all three models, the estimates of heritability decrease for both subsets of the data, although this effect is appreciable only for the IS Model. A more thorough check for consistency within the models was performed by dropping one correlation at a time. As can be seen in Table 6, the resulting heritability estimates change significantly only with the removal of the spousal correlation in the IP Model and with the removal of the correlation for monozygotic twins raised apart, r_{MZA}, in the IS Model. We shall discuss these effects in turn.

The weighted average correlation in IQ between spouses, r_M, was 0.33 in the studies summarized by Bouchard and McGue (1981), and ranged from 0.15 to 0.75. The influence on the heritability of the value of the spousal correlation was determined by varying r_M in the data and estimating the parameters of the model. Figure 2 shows the estimates of h^2 for nine values of m in the range 0.1 to 0.9 in 0.1 increments. The figure shows that removing or changing the observed spousal correlation may have a large effect on estimated heritability, since this correlation is especially important in determining the amount of resemblance between relatives that can be attributed to assortative mating.

The estimate for heritability in the IS Model is also sensitive to the correlation between monozygotic twins raised apart. This is a worrisome property since there are many reasons to suspect that data on monozygotic twins raised apart do not meet the assumptions of the models. First, these twins might be similarly influenced by intrauterine and pre-adoptive environments and by the common experience of being

adopted. Secondly, selective placement of twins is likely.

A more important aspect of the data on monozygotic twins raised apart is also apparent from Table 6. Although the fit of each model varies with each subset of data, a clearly significant improvement in fit is obtained in all three models when r_{MZA}, r_{SFF}, or r_{MMT} is removed (recall that $\chi^2_{1[0.05]} = 3.841$). In fact, the sum of squared deviations almost halves if both the monozygotic twins raised apart, MZA, and the mid-offspring-mid-parent, MMT, correlations are removed. The poor fit of r_{MMT} is probably due to our assumption that the mid-offspring value was calculated from two children per family (see Appendix B). This was due to the lack of information about the number of children in the published data, and because this assumption may very well be wrong we will omit r_{MMT} from subsequent analyses. The effect of the removal of r_{MZA} is more interesting. In all the models monozygotic twins raised apart are assumed to share no environmental or cultural similarities. Adoptive placement may not be random, however, and placement in foster homes may often occur after an extended period during which the children are together (Goldberger, 1978; Bouchard et al., 1990). The data in Table 7 are presented by Goldberger (1978) to illustrate the point that adoptive placement is not random, but rather adoptive and biological parents tend to have similar IQs and educational levels. Evidence that the adoptive homes of monozygotic twins raised apart are also correlated is shown in Table 8, from Bouchard et al. (1990). These authors argue that selective placement does not alter the correlation in IQ between monozygotic twins raised apart because the observed correlation between any one of these environmental variables and IQ is low. The combined effect of numerous correlates in the environmental and cultural experiences of adoptees may, however, be substantial in making twins raised apart fairly similar, even though no single measured correlation is very large. The high correlation between monozygotic twins raised apart may therefore be due to a correlation between their non-transmitted environmental components. It may also be due to non-additive genetic effects, namely dominance or epistasis, which may be important in monozygtic twins who share the exact same genotype not just the same alleles.

Tables 7 and 8 here.

If dominance and epistasis led to a higher correlation among monozygotic twins raised apart, we would also predict that these effects would lead to a higher than

expected correlation among monozygotic twins raised together. This additional source of similarity would be absorbed in the value of c_{mz}, which is indeed high (Table 4). However, the estimated value of c_{dz} is also high and more than double the genetically comparable value of c_s (Tables 4). Thus, dizygotic twins show a distinct "twin effect" in the environmental correlation, and it is not unreasonable to presume that this "twin effect" applies to monozygotic twins as well. This argument indicates that non-additive genetic effects probably account for only a part of the elevated r_{MZA} correlation.

We altered the three models by introducing another parameter, c_{mza}, to measure the degree of common environment in monozygotic twins raised apart, and found a dramatic decrease in the sum of squared deviations. Table 9 shows the best fitting parameter sets when c_{mza} is added to the models and r_{MMT} is deleted as a data point. The fits so obtained were very good in all three models. The estimates of c_{mza} are very high and comparable to the estimates of c_{mz} for monozygotic twins raised together. One reason that c_{mza} is very high may be that monozygotic twins raised apart share common cultural experiences as well as common non-transmitted environments. In other words, they may have similar cultural values, B or $P_M + P_F$, and not just similar environmental values, E. The cultural influences (direct or indirect) of monozygotic twins raised apart may be correlated either because of early contact with the biological parents or because of contact with relatively similar foster parents. Similar cultural and environmental experiences as well as non-additive genetic effects (dominance and epistasis) may each play a role in the high correlation observed between monozygotic twins raised apart. The fact that we estimate c_{mza} to be high is evidence that at least one of these factors is important.

Table 9 here.

Independent estimation of the effects of common environment, dominance, and epistasis is very difficult because they contribute to the correlations between relatives in much the same manner. Rao et al. (1976) argued that "environment common to relatives is likely to be a more important source of variation than is dominance," but firm conclusions cannot be drawn from the observed correlations. More information is clearly needed about the unique personal histories of separated monozygotic twins, since the standard assumptions made in models of inheritance about monozygotic

twins raised apart simply do not fit the data. The present analysis raises questions about the validity and applicability of heritability estimates based solely upon monozygotic twin data (such as that of Bouchard et al., 1990). Such treatments must ignore the contribution of non-transmitted environment, and generally provide results very different from analyses incorporating the other familial data summarized by Bouchard and McGue (1981).

The inclusion of parameters to describe the influence of non-inherited (common) environmental experiences on the correlation between siblings and twins has a profound influence on the analysis of data. When e^2 is large enough, the correlation between environmental experiences, c_i, essentially measures the extent to which the observed correlation between the siblings or twins exceeds the expected correlation, where the expected correlation is based on the parameter estimates obtained from the remaining relatives. In this sense the inclusion of the coefficient c_i drastically lowers the influence of the observed correlation r_I on the parameter estimates. If common environment were truly unimportant and if the models were accurate, then there should be little difference between the results of data analysis with and without this factor, and the estimates of $c_{mz}, c_{mza}, c_{dz}, c_s$, and c_{hs} should be small. We have found, however, that the best fitting model to the IQ data includes a large influence of common environment (Table 9). Ignoring the influence of common environmental factors by setting the coefficients, c_i, to zero, as is commonly done in the behavioral genetics literature, produces a marked increase in the estimated heritability, leads to an extremely significant drop in the goodness of fit of the models to the data, and makes the parameter surface more difficult to search for a minimal SSD (Table 10). This indicates that the IQs of siblings and twins are more similar than would be expected on the basis of the resemblance between other relatives. If environmental factors are allowed to contribute to sibling resemblance, then this excess sibling resemblance is attributed to these factors and the fit of the model improves dramatically.

Table 10 here.

The studies of IQ summarized by (Bouchard and McGue, 1981) provide the most detailed information available about the distribution of variation among relatives in a behavioral trait. To analyze this data, several measures of heritability may be employed, each providing different estimates of the importance of genetic variation as summarized in Table 11. Each heritability measure is based upon a different

set of simplifying assumptions about genetic and cultural inheritance. The lack of concordance among these estimates is a testament to the imprecision of these simple models when applied to complex traits.

Table 11 here.

4.5 Analysis of data on variation in personality characters

Eaves, Eysenck and Martin (1989) present data and analyses based upon their observations on variation in personality traits. Their working hypothesis is that there are three "super-factors" that suffice to describe major differences in personality between people. These synthetic characters are *psychoticism, extraversion*, and *neuroticism*, and the trait value of an individual is determined from a collection of directly observable, intercorrelated traits that are grouped together under these labels. Table 12 shows the correlations between various relatives for these three characters as reported by Eaves et al. (1989) (from their Table 6.3).

Table 12 here.

The three models in Table 1 may be used to analyze the data given in Table 12, and the best fitting parameter estimates are shown in Tables 13, 14 and 15. The tables record the estimates obtained in the models both with and without the parameters accounting for common environment, c_s, c_{dz}, and c_{mz}. Inclusion of the c_i's uniformly lowered estimates for heritability (h^2) and, for the psychoticism and neuroticism data, significantly improved the fit of the model (compare the difference between the χ^2 values at the bottom of each table with χ^2_3; the IS Model of the psychoticism data is only significant at the 0.1 level, the remaining comparisons are significant at the 0.01 level). Particularly striking was the reanalysis of the neuroticism data (Table 15) where the indication of appreciable genetic heritability disappeared when monozygotic twins were allowed to have similar non-transmissible environmental experiences, i. e., when c_{mz} was allowed to be non-zero. In this case, the random environmental effects, E, and the family environment, B, sufficed to explain the observed correlations between relatives. The influence of cultural transmission (as measured by $b\beta$) increased from being negligible to substantial with the inclusion of common environment parameters in the model.

Tables 13, 14 and 15 here.

Eaves et al. (1989) used a model of phenotypic transmission similar to the model of direct phenotypic transmission described in Section 1.2 in order to estimate the determinants of their three personality measures. These authors assumed, however, that while there is error in cultural transmission, there do not exist any other non-transmitted environmental influences. That is, $e = 0$ and $e' = \sqrt{b^2\delta^2}$ in the terminology of equation (12) (recall that elsewhere we drop the prime from e' for convenience). By assuming that $e = 0$, these authors were then able to estimate b, not just $b\beta_D$, from either equation (1) or equation (12) for the phenotypic variance. Eaves et al. (1989) also assumed that whatever factors cause cultural transmission to be imperfect, they are not correlated among siblings ($c_{mz} = c_{dz} = c_s = 0$). A final difference between their model and the closest of our models, the DP Model, is that Eaves et al. allowed for possible effects of dominance and epistasis by augmenting the correlation r_{MZT} by an amount Δ^2 and both r_{SST} and r_{DZT} by an amount $\frac{1}{4}\Delta^2$. Using the minimum SSD method of estimation described in Section 4.3, Eaves et al. estimated the path coefficients shown in Table 16 which refers to their Table 6.5. This table summarizes their results for both the model with the fewest parameters that still fits the data and for their full model. They claim that the full model provides a conservative estimate of heritability, and calculate two estimates of heritability: h^2 and $h^2_{broad} = h^2 + \Delta^2$. We present these estimates alongside some standard estimates of heritability in Table 17.

Table 16 and 17 here.

Our estimates of heritability (h^2) for psychoticism and extraversion obtained using the IP, DP, and IS models with the c_is were uniformly lower than the estimates of heritability considered to be the most conservative by Eaves et al. (1989). Independent of the model used, a negligible influence is ascribed to the additive effect of genes on neuroticism. The high monozygotic twin correlation observed for neuroticism is, however, attributed to an effect of dominance by Eaves et al. (1989) whereas in the models we employ this correlation is attributed to similar environmental experiences of monozygotic twins (through c_{mz}) and to similar cultural experiences ($b\beta$). As we have already argued, it is very hard to separate the effects of common environments and the effects of dominance or epistasis. Therefore, the data cannot be argued to support such claims as:

> The results ... give absolutely no indication that the environment shared by children within a family has any lasting effect on their resemblance for the two major dimensions of personality most widely studied: extraversion and neuroticism. These data, taken together with the other data that we have already described, suggest that personality theory can gain little by assuming that social learning from parents plays a major role in the creation of personality differences.
>
> — Eaves et al. (1989), p. 159.

The sensitivity of the results to particular data points was analyzed by removing one correlation at a time (as in Table 6). Heritability changes appreciably (0.05 or more) in only a few cases, and in these, the estimated value for the removed correlation (calculated from the parameter estimates obtained from an analysis of the reduced data set) is either much lower or much higher than observed, as shown in Table 18. Thus we can see, for instance, that foster sibs (adopted/adopted) had an observed correlation for neuroticism that was much higher than expected from the other relatives. We do not know, however, if these discrepancies arise from insufficient data or inaccurate models. As long as such discrepancies exist, however, there must remain considerable doubt that the models accurately describe even the major, linear determinants of behavior since the models cannot predict a removed correlation from the remaining correlations, even though the fit of the model, as measured by the SSD, may be very good (Table 15).

Table 18 here.

In this section, we have shown that estimates of heritability for the personality data are especially sensitive to whether or not the model allows for correlations between the non-transmissible environmental experiences of individuals within a family. When a correlation between the environmental values of siblings and of twins was included, the fit of the models to the data significantly improved and the heritability measures dropped substantially for the data on psychoticism and neuroticism. Therefore the full model used by Eaves et al. (1989) cannot really be said to provide a conservative estimate for heritability since the range of models analyzed by these authors was limited and did not include correlations between environmental values.

Even with two very similar linear models of inheritance, different estimates of h^2 are obtained:

Character	Eaves et al. (1989)		DP Model	
	Full model	Reduced model	With c_is	Without c_is
Psychoticism	0.22	0.36	0.16	0.35
Extraversion	0.45	0.46	0.40	0.47
Neuroticism	0.00	–	0.04	–

where only estimates for models that adequately fit the data have been included. We have used the DP Model which is the model most immediately comparable to that of Eaves et al. (1989). The reduced model of Eaves et al. (1989) and the DP Model without c_is produce nearly identical heritability estimates. These estimates drop, however, in the full models, especially in the DP Model with c_is. This table thus illustrates that even among similar linear models heritability estimates can change appreciably depending on the assumptions made about how traits are determined and what causes the resemblance of relatives.

5 Conclusions

For various reasons, research into the etiology of traits such as IQ and personality measures have focussed on heritability as a measure of the importance of genes (see, for example, Jensen, 1969). There are many different arguments against the utility of the concept of heritability in human behavioral genetics (see, for instance, Feldman and Lewontin (1975) or Lewontin et al. (1984)). That genes influence the unique composition of traits that characterize an individual is not at issue. What is at issue is whether or not we have the means to determine from a set of correlations between relatives the relative importance of genes in the etiology of a particular trait and whether a single number can encapsulate the role of genetics (or the role of culture). Simple models may fit the observed data, but they need not mirror the actual paths of influence for a trait. Indeed conceptually different *linear* models (see Section 1.4) can lead to different estimates for genetic and cultural heritability and yet *each model*

fits the data. Failing to reject a model is simply not the same as proving that it describes reality.

The analysis described above has focused on the power and sensitivity of the current methods used to determine the value of genetic and cultural heritability. For instance, by analyzing partial data sets (formed by removing a correlation from a full data set), we have attempted to determine which particular familial correlations are least consistent with the models we have used. Correction for the effect of a common rearing environment between siblings raised together removes most of the major inconsistencies between the data and the basic transmission models. The correlation between monozygous twins raised apart, however, appears to be equally inconsistent using any of the models; the fit of the models to the data improves considerably if either this correlation is removed or a shared environmental influence is allowed to act upon these twins. Additional biological resemblance due to effects of dominance and epistasis cannot be ruled out, but an elevated correlation between dizygous twins over that of ordinary full sibs points more towards a distinct twin effect. Monozygotic twins raised apart therefore are not simply related by their common genes; cultural influences and specific environmental experiences that these twins have in common might be critically important in explaining their correlation.

In all of the above models we assumed that exposure to similar non-inherited environmental factors may contribute to the similarity of siblings and twins raised together. The potential importance of correlations between the non-transmitted environmental values of individuals was pointed out by Cloninger et al. (1979a). Many authors (Bouchard et al., 1990; Eaves et al., 1989; Plomin et al., 1990), however, have assumed that no correlations exist between the specific environments of individuals. Without exception, those models that do allow for an influence of 'common environment' arrive at a lower estimate for genetic heritability. For instance, in the analysis of IQ data, heritability is often found to be around 70% when common environment is ignored but only 30 to 40% when it is included. Clearly, an estimate of heritability is only as good as our knowledge of the factors that are important in determining a trait and more attention should be paid to the details of non-transmitted similarities.

There are three main reasons why siblings (and twins) are similar. They share a portion of their genes (genetic similarity). They share parents who create and pass on a social environment (cultural similarity). Finally, they grow up together (common

environment). The difference between cultural inheritance and common environment is subtle and deserves further comment. A child's experiences are in part influenced by what the child's parents want. For instance, parents who are well educated might send their children to good schools thereby contributing to cultural transmission. But a large proportion of the child's experiences may be out of direct parental control. For instance, a parent may not be able to determine the attitude of a particular teacher towards the child. To some degree these idiosyncratic experiences are shared between siblings since sibs often attend the same schools, watch the same television programs, and live in the same neighborhood. These are elements of a common environment.

What happens when common environment is included in the study of the aggregation of traits in families? If common environment were truly unimportant (and the other assumptions of the model were appropriate) then little difference should be found between estimates with and without this factor. That is, estimates of the parameters that measure this factor should be small. We have found, however, that the best fitting models to most data sets include a large influence of common environment. The inclusion of this factor improves the fit to the data while simultaneously decreasing the estimate for genetic heritability. The basic explanation for this finding is that siblings and twins are more similar than would be expected on the basis of the resemblance of other relatives. If non-inherited factors contribute to sibling resemblance, then the excess sibling resemblance is attributable to these factors.

Intelligence and personality are too complex to know what factors really are critical during their development. We have found that heritability estimates are particularly sensitive to the inclusion or exclusion of non-inherited, common environment as a contributor to sibling resemblance. They are also sensitive to the rate of assortative mating and to the model used to analyze the data, properties which are undesirable in light of our limited knowledge about the validity of any specific model for the determination of human behavioral traits.

A Path Analysis

Path analysis is a convenient method of describing a network of linearly related variables. The basic relation in a path model is between a dependent variable Y and a series of source variables $(X_1, X_2, \ldots, X_n)$. Y is supposed to be a linear function of $X_1, X_2, \ldots, X_n$ apart from an error term which is independent of the X_i. All variables are assumed to be normally distributed and the collection of variables follows a multidimensional Gaussian distribution. For convenience problems of scaling and location of the variables are eliminated by normalization, so that the mean of every variable is zero and the variance is unity:

$$\mathrm{E}(Y) = 0, \quad \mathrm{Var}(Y) = 1, \tag{92}$$

and

$$\mathrm{E}(X_i) = 0, \quad \mathrm{Var}(X_i) = 1 \quad \text{for } i = 1, 2, \ldots, n. \tag{93}$$

The basic relation between the dependent and source variables is specified by

$$Y = \sum_{i=1}^{n} \alpha_i X_i + \epsilon W, \tag{94}$$

where the α_i measure the strength of the dependence of Y on each source variable (X_i) and W is a normalized error variable [$\mathrm{E}(W) = 0$, $\mathrm{Var}(W) = 1$ and $\mathrm{Cov}(W, X_i) = 0$ for $i = 1, 2, \ldots, n$].

Each α_i is seen as describing the determination of Y by X_i through a causal path from X_i to Y and is therefore called the *path coefficient* of that path. The path coefficients are partial regression coefficients, in that

$$\frac{\mathrm{Cov}(Y, X_i | X_j,\ j \neq i)}{\mathrm{Var}(X_i | X_j,\ j \neq i)} = \alpha_i. \tag{95}$$

Thus α_i is the regression of Y on X_i obtained when all other source variables are held constant.

Since all variables have unit variance, equation (94) entails that

$$1 = \sum_{i=1}^{n} \alpha_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_i \alpha_j r_{ij} + \epsilon^2, \tag{96}$$

where $r_{ij} = \mathrm{Cov}(X_i, X_j)$, $i,\ j = 1, 2, \ldots, n$, are the correlation coefficients between the source variables. Thus the error coefficient (ϵ) is specified completely in terms of the path coefficients (α_i) and the correlation coefficients (r_{ij}). This specification may then be used in equation (94), so that Y is determined completely in terms of the α_i and r_{ij}.

A.1 Covariance calculations

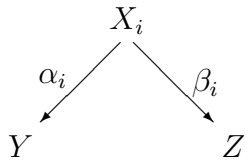
Consider two variables Y and Z (both normalized) with the (normalized) sources $X_1, X_2, \ldots, X_n$, where Y and Z are described by the path coefficients α_i and β_i, $i = 1, 2, \ldots, n$, respectively:

$$\begin{aligned} Y &= \sum_{i=1}^{n} \alpha_i X_i + \epsilon_Y W_Y, \\ Z &= \sum_{i=1}^{n} \beta_i X_i + \epsilon_Z W_Z. \end{aligned}$$

Then the correlation between these two variables is

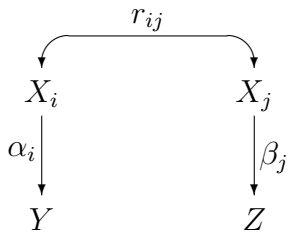
$$\mathrm{Cov}(Y, Z) = \sum_{i=1}^{n} \alpha_i \beta_i + \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_i \beta_j r_{ij}, \tag{97}$$

since W_Y and W_Z are assumed to be independent. This correlation arises when single source variables (e. g. the source X_i) have an influence on both dependent variables:



(98)

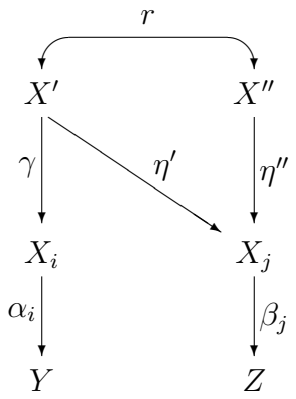
and when correlated source variables (e. g. the sources X_i and X_j) have separate influences on the two dependent variables:



(99)

The correlation between Y and Z is then the sum of the correlations from all common sources and from all correlated sources.

This result is the foundation of path analysis and it may be used recursively in a network of interrelated variables. If the correlation between, e. g., X_i and X_j is further specified in terms of their common and correlated sources, then r_{ij} may be specified by an equation analogous to (97). By this operation the correlation between Y and Z in (99) is specified in terms of correlated and common sources of X_i and X_j:



(100)

Equation (97) then becomes a sum of terms like $\alpha_i\gamma\eta'\beta_j$ and $\alpha_i\gamma r\eta''\beta_j$. The correlation between the two variables Y and Z is therefore obtained as a sum of contributions from common ultimate sources and correlated ultimate sources, where we call a source that is not specified by previous source variables an ultimate source. Each contribution is a product of the path coefficients along a path leading back from one variable to one ultimate source and the product of those coefficients along a path leading forward from a second ultimate source to the other variable, including a factor which is the correlation between the two ultimate sources (which takes the value one in the case

of a common source).

Thus, the correlation between any two variables (such as Y and Z) is simply the sum of the paths connecting the two variables when the following rules are observed:

1. A path connecting two variables must must move up through the path diagram and then back down (up to a common cause or a pair of correlated causes and then back down).

2. A path may consist of any number of segments described by path coefficients, but only one correlation may be used in the contribution from any one path *and* that correlation must be at the top of the path between the ultimate causes of the path.

3. An ultimate cause is the only cause that may appear in both branches of a path. The path up to the common cause or pair of correlated causes and the path back down cannot pass through the same individual (except at the top of the path).

In the text, we used a covariance decomposition method to calculate the expected correlation between a pair of relatives. This method expanded a given correlation using appropriate equations for phenotypic determination, genetic inheritance, and cultural inheritance until variables were obtained whose correlations were known. Covariance decomposition and path analysis are completely equivalent, although we have found the former to be less susceptible to errors in record-keeping.

A.2 Sample Path Diagrams

A path diagram may be constructed that specifies all the paths that contribute to the correlation between any two relatives. For the indirect transmission model (B to B inheritance), we illustrate three path diagrams: one connecting a mother and her offspring (Fig. 3), one between a foster parent and its offspring (Fig. 4), and one between a biological parent and its offspring who was adopted away at birth (Fig. 5). The correlations between the latent values of the mates ($r_{AA}, r_{AB}, r_{AE}, ...$) are specified by the covariance matrix $\mathbf{\Theta}$, given by equation (32) for phenotypic assortative mating and equation (37) for social assortative mating.

B Familial Correlations

B.1 Definitions

We use the following symbols to denote the correlation between the phenotypic values of a pair of relatives:

r_{MZT} = Monozygotic twins raised together
r_{MZA} = Monozygotic twins raised apart (neither in the biological home)
r_{DZT} = Dizygotic twins raised together
r_{DZA} = Dizygotic twins raised apart (neither in the biological home)
r_{SST} = Full sibs raised together
r_{SSA} = Full sibs raised apart (neither in the biological home)
r_{SFF} = Two unrelated individuals raised in the same home (both are adopted)
r_{SFB} = Two unrelated individuals raised in the same home (only one is adopted)
r_{OPT} = A parent and his or her biological child (together)
r_{OMT} = Mid-parent and their biological child (together)
r_{MMT} = Mid-parent and mid-offspring (together)
r_{OPA} = A parent and his or her biological child (separated)
r_{FOP} = A parent and his or her foster child
r_{FMP} = Mid-parent and their foster child
r_{HSS} = Half-sibs, both raised by their respective biological parents (no cultural influence of the step-parents)
r_{ANT} = Aunt/uncle and niece/nephew (no separations)
r_{CZ1} = First cousins (no separations)
$r_{C_{ij}}$ = The i^{th} descendent of one sib and the j^{th} descendent of the other sib (no separations)
r_{GPO} = Grandparent and grandchild (no separations)
r_{V_n} = An individual and his/her n^{th} descendent (no separations)

Two comments about the familial correlations are in order. Correlation r_{MMT} is a measure that depends on the number of offspring used to calculate the mid-offspring phenotypic value. We arbitrarily assume that the mid-offspring value is an average of the phenotypic values of exactly two offspring. For r_{HSS}, we assume that half-sibs are raised separately by their biological parents in the two families that each

contain the common biological parent. Several other conditions of rearing are possible for half-sibs, although published data often lump all types of half-siblings together. Cloninger et al. (1979b) discuss different circumstances of rearing for half-sibs and provide correlations for their model. Assuming different half-sib relations (e.g. half-sibs that are raised together by one of the sets of parents, or half-sibs that are both raised by the single common parent) in our analysis of IQ data (section 4.4) does not appreciably change the heritability estimate or the sum of squared deviations, but it does alter the estimated correlation between non-transmitted environmental values of half-sibs (c_{hs}).

B.1.1 Partial regression coefficients among remote relatives.

The calculation of the correlations between remote relatives is simplified by introducing the functions K_n, L_n and M_n which are defined in the path analyses of Rao et al. (1979) and Cloninger et al. (1979a). We also introduce the function, N_n, which is related to, but of more general utility than M_n. The function K_n is the partial regression coefficient describing the dependence of the phenotype of an $(n-1)^{th}$ generation descendent, P_{n-1}, on the genotype of an individual *given the cultural value of the individual*, that is

$$K_n = \frac{\mathrm{Cov}(P_{n-1}, A|B, E)}{\mathrm{Var}(A|B, E)}.$$

For instance, if $n = 1$ then $K_1 = h$ from equation (4) because P_0 is the phenotype of the considered individual. If $n = 2$ then

$$K_2 = \frac{\mathrm{Cov}(P_1, A|B, E)}{\mathrm{Var}(A|B, E)} = \frac{\mathrm{Cov}(P_O, A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)}$$

in our usual notation. In the case of indirect cultural transmission (equation (8)), we have

$$\begin{aligned} K_2 &= \frac{\mathrm{Cov}(P_O, A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)} = \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)} \\ &= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_I(B_M + B_F), A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)} \\ &= \tfrac{1}{2}h\frac{\mathrm{Cov}(A_M, A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)} + \tfrac{1}{2}h\frac{\mathrm{Cov}(A_F, A_M|B_M, E_M)}{\mathrm{Var}(A_M|B_M, E_M)} \end{aligned}$$

$$
\begin{aligned}
& + b\beta_I \frac{\mathrm{Cov}(B_F, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} \\
= \; & \tfrac{1}{2}h(1 + c_{AA}) + b\beta_I c_{BA},
\end{aligned}
$$

where the partial regression coefficients c_{AA} and c_{BA} are calculated in section 1.3.3. In the case of direct cultural transmission (equation (10)), we have

$$
\begin{aligned}
K_2 &= \frac{\mathrm{Cov}(P_O, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} = \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} \\
&= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_D(P_M + P_F), A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} \\
&= \tfrac{1}{2}h \frac{\mathrm{Cov}(A_M, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} + \tfrac{1}{2}h \frac{\mathrm{Cov}(A_F, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} \\
&\quad + b\beta_D \frac{\mathrm{Cov}(hA_M + bB_M + eE_M + hA_F + bB_F + eE_F, A_M | B_M, E_M)}{\mathrm{Var}(A_M | B_M, E_M)} \\
&= \tfrac{1}{2}h(1 + c_{AA}) + b\beta_D(h + hc_{AA} + bc_{BA} + ec_{EA}).
\end{aligned}
$$

Similarly, L_n is defined as the regression coefficient of the phenotype of an $(n-1)^{th}$ descendent on the cultural value of an individual *given the genotypic value of the individual*, that is

$$
L_n = \frac{\mathrm{Cov}(P_{n-1}, B | A, E)}{\mathrm{Var}(B | A, E)}.
$$

Thus, $L_1 = b$ from equation (4), and

$$
L_2 = \frac{\mathrm{Cov}(P_1, B | A, E)}{\mathrm{Var}(B | A, E)} = \frac{\mathrm{Cov}(P_O, B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)}
$$

in our usual notation. With indirect cultural transmission we have

$$
\begin{aligned}
L_2 &= \frac{\mathrm{Cov}(P_O, B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)} = \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)} \\
&= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_I(B_M + B_F), B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)} \\
&= \tfrac{1}{2}h \frac{\mathrm{Cov}(A_F, B_M | B_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)} + b\beta_I \frac{\mathrm{Cov}(B_M, B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)} \\
&\quad + b\beta_I \frac{\mathrm{Cov}(B_F, B_M | A_M, E_M)}{\mathrm{Var}(B_M | A_M, E_M)}
\end{aligned}
$$

$$= \frac{1}{2}hc_{AB} + b\beta_I(1 + c_{BB}),$$

and with direct cultural transmission we have

$$\begin{aligned}
L_2 &= \frac{\mathrm{Cov}(P_O, B_M|A_M, E_M)}{\mathrm{Var}(B_M|A_M, E_M)} = \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, B_M|A_M, E_M)}{\mathrm{Var}(B_M|A_M, E_M)} \\
&= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_D(P_M + P_F), B_M|A_M, E_M)}{\mathrm{Var}(B_M|A_M, E_M)} \\
&= \frac{1}{2}h\frac{\mathrm{Cov}(A_F, B_M|A_M, E_M)}{\mathrm{Var}(B_M|A_M, E_M)} \\
&\quad + b\beta_D\frac{\mathrm{Cov}(hA_M + bB_M + eE_M + hA_F + bB_F + eE_F, B_M|A_M, E_M)}{\mathrm{Var}(B_M|A_M, E_M)} \\
&= \frac{1}{2}hc_{AB} + b\beta_D(b + hc_{AB} + bc_{BB} + ec_{EB}).
\end{aligned}$$

The function M_n is the partial regression coefficient of the phenotype of an $(n-1)^{th}$ descendent on the phenotype of an individual *given complete information about the individual's mate*, that is

$$M_n = \frac{\mathrm{Cov}(P_{n-1}, P|A_{spouse}, B_{spouse}, E_{spouse})}{\mathrm{Var}(P|A_{spouse}, B_{spouse}, E_{spouse})}.$$

Here $M_1 = 1$, and M_2 is the regression coefficient

$$\frac{\mathrm{Cov}(P_O, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}.$$

In the model of indirect cultural inheritance, this becomes:

$$\begin{aligned}
M_2 &= \frac{\mathrm{Cov}(P_O, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)} \\
&= \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)} \\
&= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_I(B_M + B_F), P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)} \\
&= \frac{\mathrm{Cov}(\frac{1}{2}hA_M + b\beta_I B_M, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}.
\end{aligned}$$

With phenotypic homogamy but *not with social homogamy*, we have that

$$\begin{aligned}\mathrm{Cov}(A_M, P_M|A_F, B_F, E_F) &= \mathrm{Cov}(A_M, P_M|P_F),\\ \mathrm{Cov}(B_M, P_M|A_F, B_F, E_F) &= \mathrm{Cov}(B_M, P_M|P_F),\end{aligned}$$

and these conditional covariances are independent of the particular value of P_F in the multivariate Gaussian distribution (Graybill, 1961, Theorem 3.10). Using this fact and the conditional expansion of a covariance, we obtain

$$\begin{aligned}\mathrm{Cov}(A_M, P_M) &= \mathrm{E}\Big(\mathrm{Cov}(A_M, P_M|P_F)\Big) + \mathrm{Cov}\Big(\mathrm{E}(A_M|P_F), \mathrm{E}(P_M|P_F)\Big)\\ &= \mathrm{Cov}(A_M, P_M|P_F) + \mathrm{Cov}(m\gamma P_F, mP_F).\end{aligned}$$

By this line of proof, we can show that

$$\begin{aligned}\mathrm{Cov}(A_M, P_M|A_F, B_F, E_F) &= \gamma(1-m^2),\\ \mathrm{Cov}(B_M, P_M|A_F, B_F, E_F) &= \phi(1-m^2),\\ \mathrm{Var}(P_M|A_F, B_F, E_F) &= (1-m^2).\end{aligned}$$

For the model of indirect cultural transmission with phenotypic homogamy this produces

$$M_2 = \tfrac{1}{2}h\gamma + b\beta_I\phi$$

(Cloninger et al., 1979a). In the model of direct cultural transmission with phenotypic homogamy, we have

$$\begin{aligned}M_2 &= \frac{\mathrm{Cov}(P_O, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}\\ &= \frac{\mathrm{Cov}(hA_O + bB_O + eE_O, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}\\ &= \frac{\mathrm{Cov}(\frac{1}{2}h(A_M + A_F) + b\beta_D(P_M + P_F), P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}\\ &= \frac{\mathrm{Cov}(\frac{1}{2}hA_M + b\beta_D P_M, P_M|A_F, B_F, E_F)}{\mathrm{Var}(P_M|A_F, B_F, E_F)}\\ &= \tfrac{1}{2}h\gamma + b\beta_D.\end{aligned}$$

Finally, we introduce the function, N_n, which may replace M_n in the calculation of the correlations between relatives, and which does not depend on the mode of assortment. The function N_n is the partial regression coefficient describing the dependence of the phenotype of an $(n-1)^{th}$ descendent on the environmental value (E) of an individual:

$$N_n = \frac{\mathrm{Cov}(P_{n-1}, E)}{\mathrm{Var}(E)} = \mathrm{Cov}(P_{n-1}, E).$$

N_n is not conditioned on the genotypic and cultural values of the ancestral individual since by assumption these values do not affect her or his specific environmental value. Then $N_1 = e$ from equation (4) and $N_2 = \mathrm{Cov}(P_1, E) = \mathrm{Cov}(P_O, E_M)$. For the model of indirect cultural transmission it is

$$\begin{aligned} N_2 &= \mathrm{Cov}(P_O, E_M) = \mathrm{Cov}(hA_O + bB_O + eE_O, E_M) \\ &= \mathrm{Cov}(\tfrac{1}{2}h(A_M + A_F) + b\beta_I(B_M + B_F), E_M) \\ &= \tfrac{1}{2}hr_{AE} + b\beta_I r_{BE}, \end{aligned}$$

and for the model of direct cultural transmission we obtain

$$\begin{aligned} N_2 &= \mathrm{Cov}(P_O, E_M) = \mathrm{Cov}(hA_O + bB_O + eE_O, E_M) \\ &= \mathrm{Cov}(\tfrac{1}{2}h(A_M + A_F) + b\beta_D(P_M + P_F), E_M) \\ &= \mathrm{Cov}(\tfrac{1}{2}hA_F + b\beta_D(hA_M + bB_M + eE_M + hA_F + bB_F + eE_F), E_M) \\ &= \tfrac{1}{2}hr_{AE} + b\beta_D(e + hr_{AE} + br_{BE} + er_{EE}). \end{aligned}$$

For the social homogamy model described by equation (37), r_{AE}, r_{BE}, and r_{EE} equal zero and the N_n become zero for $n > 1$ with indirect cultural inheritance.

N_n measures the influence of a specific environmental value (E) on the $(n-1)^{th}$ descendent of an individual. This function will be important in any pedigree in which the environmental values of individuals are correlated. For instance, the environmental values of siblings are correlated by an amount c_s and this correlation will add $N_2^2 c_s$ to the correlation between first cousins. With indirect cultural transmission and phenotypic homogamy, it can be shown that $N_n = meM_n$ for all $n > 1$ (M_1 is undefined) and either measure may be used. With direct cultural transmission or with social homogamy, however, N_n and M_n are not related in a simple way and the correlations between relatives are composed of terms that include N_n (not M_n).

B.2 Indirect cultural transmission and phenotypic homogamy

$$
\begin{aligned}
r_{MZT}{}^{a} &= h^2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_{mz}e^2 \\
r_{MZA}{}^{a} &= (h^2 + c_{mza}e^2)/(1-2whb) \\
r_{DZT}{}^{a} &= h^2(1+r_{AA})/2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_{dz}e^2 \\
r_{DZA} &= h^2(1+r_{AA})/[2(1-2whb)] \\
r_{SST}{}^{a} &= h^2(1+r_{AA})/2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_s e^2 \\
r_{SSA} &= h^2(1+r_{AA})/[2(1-2whb)] \\
r_{SFF}{}^{a} &= [2\beta_I^2(1+r_{BB})b^2 + c_s e^2]/(1-2whb) \\
r_{SFB}{}^{a} &= [2\beta_I^2(1+r_{BB})b^2 + whb + c_s e^2]/\sqrt{1-2whb} \\
r_{OPT}{}^{a} &= (1+m)[h(h+wb)/2 + \beta_I b(b+wh)] \\
r_{OMT} &= \sqrt{2/(1+m)}\, r_{OPT} \\
r_{MMT} &= \sqrt{2/(1+m)}\sqrt{2/(1+r_{SST})}\, r_{OPT} \qquad \text{[With 2 children per family]} \\
r_{OPA} &= [(1+m)h(h+wb)/2]/\sqrt{1-2whb} \\
r_{FOP}{}^{a} &= [(1+m)\beta_I b(b+wh)]/\sqrt{1-2whb} \\
r_{FMP} &= \sqrt{2/(1+m)}\, r_{FOP} \\
r_{HSS}{}^{b} &= h^2(1+2r_{AA}+mr_{AA})/4 + \beta_I^2(1+2r_{BB}+mr_{BB})b^2 \\
&\quad + hb\beta_I(w+2r_{AB}+mr_{AB}) + c_{hs}e^2 \\
r_{ANT}{}^{a} &= h/2(1+r_{AA})K_2 + 2\beta_I^2(1+r_{BB})bL_2 \\
&\quad + \beta_I(w+r_{AB})(hL_2+bK_2) + c_s N_1 N_2 \\
r_{CZ1}{}^{a} &= 1/2(1+r_{AA})K_2^2 + 2\beta_I^2(1+r_{BB})L_2^2 \\
&\quad + 2\beta_I(w+r_{AB})L_2K_2 + c_s N_2^2 \\
r_{C_{ij}}{}^{a} &= 1/2(1+r_{AA})K_{i+1}K_{j+1} + 2\beta_I^2(1+r_{BB})L_{i+1}L_{j+1} \\
&\quad + \beta_I(w+r_{AB})(K_{i+1}L_{j+1} + L_{i+1}K_{j+1}) + c_s N_{i+1}N_{j+1} \\
r_{GPO}{}^{a} &= 1/2(1+m)(h+wb)K_2 + \beta_I(1+m)(b+wh)L_2 \\
r_{V_n}{}^{a} &= 1/2(1+m)(h+wb)K_n + \beta_I(1+m)(b+wh)L_n
\end{aligned}
$$

where

$$
\begin{aligned}
K_1 &= h \\
L_1 &= b \\
N_1 &= e \\
K_2{}^{a} &= h[1+m(h^2+whb)]/2 + b\beta_I m(hb+wh^2) \\
L_2{}^{a} &= b\beta_I[1+m(b^2+whb)] + hm(hb+wb^2)/2 \\
N_2 &= meh(h+wb)/2 + meb\beta_I(b+wh)
\end{aligned}
$$

K_n[a] $= K_{n-1}[1 + m(h^2 + whb)]/2 + L_{n-1}\beta_I m(hb + wh^2)$
L_n[a] $= L_{n-1}\beta_I[1 + m(b^2 + whb)] + K_{n-1}m(hb + wb^2)/2$
$N_n \ = meK_{n-1}(h + wb)/2 + meL_{n-1}\beta_I(b + wh).$

[a] From Cloninger, Rice and Reich (1979a).
[b] From Cloninger, Rice and Reich (1979b).

The parameters estimated in this model are $h, b, \beta_I, m, c_s, c_{dz}, c_{mz}, c_{mza}$, and c_{hs} (c_{mza} is set to zero unless explicitly mentioned), which are all constrained to lie between zero and one. The correlations between the latent variables of mates (r_{AA}, r_{AB}, r_{AE},...) are derived according to equations (25)–(30). The parameters e and w are defined by equations (2) and (60):

$$\begin{aligned} 1 &= h^2 + b^2 + e^2 + 2bhw, \\ 0 &= w^2 hbm\beta_I + w(\beta_I + \beta_I mb^2 + \beta_I mh^2 - 1) + hbm\beta_I. \end{aligned}$$

The parameter estimates must satisfy two inequalities that are constraints on the estimation. First the convergence of w to the equilibrium value requires

$$1 \ > \ \beta_I \left(1 + m(h + b)^2\right)$$

from condition (61), and consistency of the parameters requires that

$$\beta_I \ < \ 1/\sqrt{2(1 + r_{BB})}$$

for δ to be real after normalization of equation (8).

B.3 Direct cultural transmission and phenotypic homogamy

$$
\begin{aligned}
r_{MZT}{}^{c} &= h^2 + 2(1+m)b^2\beta_D^2 + 2\gamma hb\beta_D(1+m) + c_{mz}e^2 \\
r_{MZA} &= (h^2 + c_{mza}e^2)/[1 - 2\gamma hb\beta_D(1+m)] \\
r_{DZT}{}^{c} &= h^2(1+m\gamma^2)/2 + 2(1+m)b^2\beta_D^2 + 2\gamma hb\beta_D(1+m) + c_{dz}e^2 \\
r_{DZA} &= h^2(1+m\gamma^2)/[2(1 - 2\gamma hb\beta_D(1+m))] \\
r_{SST} &= h^2(1+m\gamma^2)/2 + 2(1+m)b^2\beta_D^2 + 2\gamma hb\beta_D(1+m) + c_s e^2 \\
r_{SSA} &= h^2(1+m\gamma^2)/[2(1 - 2\gamma hb\beta_D(1+m))] \\
r_{SFF} &= (2(1+m)b^2\beta_D^2 + c_s e^2)/[1 - 2\gamma hb\beta_D(1+m)] \\
r_{SFB} &= (2(1+m)b^2\beta_D^2 + \gamma hb\beta_D(1+m) + c_s e^2)/\sqrt{1 - 2\gamma hb\beta_D(1+m)} \\
r_{OPT} &= (1+m)(h\gamma/2 + b\beta_D) \\
r_{OMT} &= \sqrt{2/(1+m)}\, r_{OPT} \\
r_{MMT} &= \sqrt{2/(1+m)}\sqrt{2/(1+r_{SST})}\, r_{OPT} \qquad \text{[With 2 children per family]} \\
r_{OPA} &= (1+m)(h\gamma/2)/\sqrt{1 - 2\gamma hb\beta_D(1+m)} \\
r_{FOP} &= (1+m)b\beta_D/\sqrt{1 - 2\gamma hb\beta_D(1+m)} \\
r_{FMP} &= \sqrt{2/(1+m)})\, r_{FOP} \\
r_{HSS} &= \tfrac{1}{4}h^2(1 + 2\gamma^2 m + \gamma^2 m^2) + b^2\beta_D^2(1+m)^2 + \gamma b\beta_D h(1+m)^2 + c_{hs}e^2 \\
r_{ANT} &= \tfrac{1}{2}h(1+\gamma^2 m)K_2 + 2(1+m)b\beta_D^2 L_2 + \beta_D\gamma(1+m)(hL_2 + bK_2) + c_s N_1 N_2 \\
r_{CZ1} &= \tfrac{1}{2}(1+\gamma^2 m)K_2^2 + 2(1+m)\beta_D^2 L_2^2 + 2\beta_D\gamma(1+m)(L_2K_2) + c_s N_2^2 \\
r_{C_{ij}} &= \tfrac{1}{2}(1+\gamma^2 m)K_{i+1}K_{j+1} + 2(1+m)\beta_D^2 L_{i+1}L_{j+1} \\
&\quad + \beta_D\gamma(1+m)(K_{i+1}L_{j+1} + L_{i+1}K_{j+1}) + c_s N_{i+1}N_{j+1} \\
r_{GPO} &= \tfrac{1}{2}(1+m)\gamma K_2 + \beta_D(1+m)L_2 \\
r_{V_n} &= \tfrac{1}{2}(1+m)\gamma K_n + \beta_D(1+m)L_n
\end{aligned}
$$

where

$$
\begin{aligned}
K_1 &= h \\
L_1 &= b \\
N_1 &= e \\
K_2 &= \tfrac{1}{2}h(1+mh\gamma) + (1+m)b\beta_D h \\
L_2 &= b^2\beta_D(1+m) + \tfrac{1}{2}mhb\gamma \\
N_2 &= \tfrac{1}{2}hme\gamma + b\beta_D e(1+m) \\
K_n &= K_{n-1}\tfrac{1}{2}(1+mh\gamma) + L_{n-1}\beta_D(1+m)h \\
L_n &= L_{n-1}b\beta_D(1+m) + K_{n-1}\tfrac{1}{2}mb\gamma \\
N_n &= K_{n-1}\tfrac{1}{2}me\gamma + L_{n-1}\beta_D e(1+m)
\end{aligned}
$$

[c] Consistent with Martin et al. (1986).

The parameters estimated in this model are $h, b\beta_D, m, c_s, c_{dz}, c_{mz}, c_{mza}$, and c_{hs} (c_{mza} is non-zero only when explicitly mentioned), which are all constrained to lie between zero and one. We have assumed that equation (12) rather than equation (1) describes phenotypic determination, so that any correlation between the environmental values of relatives is actually a correlation between their E' values, as defined in equation (12). For clarity, however, we have suppressed the prime notation in the above formulae (e' is given as e). In this case, b and β_D cannot be independently estimated, but their product can. Hence, in each correlation provided above, b and β_D are always paired (note that this means that L_n is always multiplied by β_D). The parameter e is estimated from equation 13,

$$1 \;=\; h^2 + 2b^2\beta_D^2(1+m) + e^2 + 2b\beta_D h\gamma(1+m),$$

and γ is defined by its equilibrium value from equation (67):

$$\gamma \;=\; \frac{h}{1 - b\beta_D(1+m)}.$$

For δ to be real after normalization of equation (10), the parameters are subject to the constraint that

$$\beta_D \;<\; 1/\sqrt{2(1+m)}.$$

B.4 Indirect cultural transmission and social homogamy

$$
\begin{aligned}
r_{MZT}{}^{d} &= h^2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_{mz}e^2 \\
r_{MZA} &= (h^2 + c_{mza}e^2)/(1-2whb) \\
r_{DZT} &= h^2(1+r_{AA})/2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_{dz}e^2 \\
r_{DZA} &= h^2(1+r_{AA})/[2(1-2whb)] \\
r_{SST} &= h^2(1+r_{AA})/2 + 2\beta_I^2(1+r_{BB})b^2 + 2whb + c_s e^2 \\
r_{SSA}{}^{d} &= h^2(1+r_{AA})/[2(1-2whb)] \\
r_{SFF}{}^{d} &= [2\beta_I^2(1+r_{BB})b^2 + c_s e^2]/(1-2whb) \\
r_{SFB} &= [2\beta_I^2(1+r_{BB})b^2 + whb + c_s e^2]/\sqrt{1-2whb} \\
r_{OPT} &= (1+r_{AA})h^2/2 + \beta_I b^2(1+r_{BB}) + hb(\beta_I + 1/2)(w + r_{AB}) \\
r_{OMT} &= \sqrt{2/(1+m)}\, r_{OPT} \\
r_{MMT} &= \sqrt{2/(1+m)}\sqrt{2/(1+r_{SST})}\, r_{OPT} \qquad \text{[With 2 children per family]} \\
r_{OPA}{}^{d} &= [(1+r_{AA})h^2/2 + hb(w+r_{AB})/2]/\sqrt{1-2whb} \\
r_{FOP}{}^{d} &= [\beta_I b^2(1+r_{BB}) + hb\beta_I(w+r_{AB})]/\sqrt{1-2whb} \\
r_{FMP} &= \sqrt{2/(1+m)}\, r_{FOP} \\
r_{HSS}{}^{d} &= h^2(1+3r_{AA})/4 + b^2\beta_I^2(1+3r_{BB}) + bh\beta_I(w+3r_{AB}) + c_{hs}e^2 \\
r_{ANT}{}^{d} &= h(1+r_{AA})K_2/2 + 2b\beta_I^2(1+r_{BB})L_2 + w(hL_2 + bK_2) \\
r_{CZ1}{}^{d} &= (1+r_{AA})K_2^2/2 + 2\beta_I^2(1+r_{BB})L_2^2 + 2wL_2K_2 \\
r_{C_{ij}} &= (1+r_{AA})K_{i+1}K_{j+1}/2 + 2\beta_I^2(1+r_{BB})L_{i+1}L_{j+1} + w(K_{i+1}L_{j+1} + L_{i+1}K_{j+1}) \\
r_{GPO}{}^{d} &= [h(1+r_{AA})/2 + b(w+r_{AB})/2]K_2 + [b\beta_I(1+r_{BB}) + h\beta_I(w+r_{AB})]L_2 \\
r_{V_n} &= [h(1+r_{AA})/2 + b(w+r_{AB})/2]K_n + [b\beta_I(1+r_{BB}) + h\beta_I(w+r_{AB})]L_n
\end{aligned}
$$

where

$$
\begin{aligned}
K_2{}^{d} &= \tfrac{1}{2}h[1 + (r_{AA} - wr_{AB})/(1-w^2)] + b\beta_I[(r_{AB} - wr_{BB})/(1-w^2)] \\
L_2{}^{d} &= b\beta_I[1 + (r_{BB} - wr_{AB})/(1-w^2)] + \tfrac{1}{2}h[(r_{AB} - wr_{AA})/(1-w^2)] \\
K_n{}^{d} &= \tfrac{1}{2}K_{n-1}[1 + (r_{AA} - wr_{AB})/(1-w^2)] + L_{n-1}\beta_I[(r_{AB} - wr_{BB})/(1-w^2)] \\
L_n{}^{d} &= L_{n-1}\beta_I[1 + (r_{BB} - wr_{AB})/(1-w^2)] + \tfrac{1}{2}K_{n-1}[(r_{AB} - wr_{AA})/(1-w^2)] \\
N_n &= 0
\end{aligned}
$$

[d] Given in a slightly different form in Rao, Morton and Cloninger (1979).

The parameters estimated in this model are $h, b, \beta_I, r_{AA}, r_{BB}, c_s, c_{dz}, c_{mz}, c_{mza}$, and c_{hs} (c_{mza} is non-zero only when explicitly included), which are all constrained to lie between zero and one. For half-siblings, we follow Rao et al. (1979) and make the assumption that spouses of the common parent are as correlated to one another as they are to the common spouse. This reflects an assumption that mates are chosen randomly from within a social grouping. Other formulations are possible and will have an effect on the estimate of c_{hs}.

The parameter r_{AB} is specified as

$$r_{AB} = \sqrt{r_{AA} r_{BB}},$$

as assumed by Rao et al. (1979). The parameters e and w are defined by equations (2) and (62):

$$\begin{aligned} 1 &= h^2 + b^2 + e^2 + 2bhw, \\ w &= \frac{\beta_I \sqrt{r_{AA} r_{BB}}}{1 - \beta_I}. \end{aligned}$$

Consistency of the parameter estimates requires that the following two inequalities be satisfied:

$$1 \geq h^2 r_{AA} + b^2 r_{BB} + 2bh r_{AB} = m,$$

for the correlation between mates, m, to be less than one, and

$$\beta_I < 1/\sqrt{2(1 + r_{BB})}.$$

for δ be real in equation (8).

C Conditional Heritabilities

Using our model (1) of phenotypic determination with genetic transmission, equation (6), and indirect cultural inheritance, equation (8), we can rewrite equation (86) using the relations

$$\begin{aligned}\mathrm{Cov}(P_M, P_O|B_M, B_O) &= \mathrm{Cov}\left[P_M, \tfrac{1}{2}h(A_M + A_F)|B_M, B_O\right] \\ &= \tfrac{1}{2}h^2\left[\mathrm{Var}(A_M|B_M, B_F) + \mathrm{Cov}(A_M, A_F|B_M, B_F)\right] \\ &\quad +\tfrac{1}{2}he\left[\mathrm{Cov}(E_M, A_M|B_M, B_F) + \mathrm{Cov}(E_M, A_F|B_M, B_F)\right]\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}(P_M|B_M, B_O) &= \mathrm{Var}\left[hA_M + eE_M|B_M, B_O\right] \\ &= h^2\mathrm{Var}(A_M|B_M, B_F) + e^2\mathrm{Var}(E_M|B_M, B_F) \\ &\quad +2eh\mathrm{Cov}(E_M, A_M|B_M, B_F).\end{aligned}$$

The variables A_M and E_M are independent even when conditioned upon the cultural values B_M and B_F and we may continue to use equations (38) and (41) to specify E_M in terms of the hidden variables in the father. The conditional convariance and variance thus become

$$\begin{aligned}\mathrm{Cov}(P_M, P_O|B_M, B_O) &= \tfrac{1}{2}h^2\left(\mathrm{Var}(A_M|B_M, B_F) + \mathrm{Cov}(A_M, A_F|B_M, B_F)\right) \\ &\quad +\tfrac{1}{2}h^2e^2m\mathrm{Var}(A_F|B_M, B_F), \qquad (101)\end{aligned}$$

$$\mathrm{Var}(P_M|B_M, B_O) = h^2\mathrm{Var}(A_M|B_M, B_F) + e^2 \qquad (102)$$

for the phenotypic homogamy model. Thus, we need to find the conditional variance-covariance matrix of (A_M, A_F) given (B_M, B_F).

The variance-covariance matrix of (A_M, A_F, B_M, B_F) may be written as

$$\mathbf{\Gamma_{AB}} = \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C} & \mathbf{B} \end{pmatrix},$$

where the elements are

$$\mathbf{A} = \begin{pmatrix} 1 & m\gamma^2 \\ m\gamma^2 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & m\phi^2 \\ m\phi^2 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} w & m\gamma\phi \\ m\gamma\phi & w \end{pmatrix}.$$

As these three matrices commute it is easy to prove that the inverse of the variance-covariance matrix is given by

$$\mathbf{\Gamma_{AB}^{-1}} = \begin{pmatrix} \mathbf{B}(\mathbf{AB}-\mathbf{C}^2)^{-1} & -\mathbf{C}(\mathbf{AB}-\mathbf{C}^2)^{-1} \\ -\mathbf{C}(\mathbf{AB}-\mathbf{C}^2)^{-1} & \mathbf{A}(\mathbf{AB}-\mathbf{C}^2)^{-1} \end{pmatrix}.$$

In the multivariate Gaussian case, the conditional variance-covariance matrix of (A_M, A_F) given (B_M, B_F) is then given by

$$\Big(\mathbf{B}(\mathbf{AB}-\mathbf{C}^2)^{-1}\Big)^{-1} = \mathbf{A} - \mathbf{C}^2\mathbf{B}^{-1} \tag{103}$$

(Graybill, 1961), and we have

$$\mathbf{B}^{-1} = \frac{1}{1-m^2\phi^4} \begin{pmatrix} 1 & -m\phi^2 \\ -m\phi^2 & 1 \end{pmatrix}.$$

The conditional parent-offspring correlation with cultural values specified, $b_{OP|\text{culture}}$, can now be rewritten as

$$\frac{\frac{1}{2}h^2\Big(1+m\gamma^2-\frac{(w+m\gamma\phi)^2}{1+m\phi^2}\Big)+\frac{1}{2}h^2e^2m\Big(1-\frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma\phi^3w}{1-m^2\phi^4}\Big)}{h^2\Big(1-\frac{w^2+m^2\gamma^2\phi^2-2m^2\gamma\phi^3w}{1-m^2\phi^4}\Big)+e^2}$$

which is equation (88). Equation (89) is obtained by a symmetry argument.

Table 1: Three linear models of inheritance.

MODEL	CULTURAL TRANSMISSION	HOMOGAMY
IP	Indirect	Phenotypic
IS	Indirect	Social
DP	Direct	Phenotypic

Table 2: Dependence of fit on starting position in the IP Model. The search begins from values specified as "starting" and ends at the values indicated as "final."

	Starting	Final	Starting	Final	Starting	Final
h	0.01	0.57	0.25	0.57	0.90	0.57
b	0.01	0.51	0.25	0.51	0.01	0.52
β	0.01	0.48	0.25	0.48	0.01	0.47
m	0.01	0.32	0.25	0.32	0.01	0.32
c_s	0.01	0.25	0.25	0.25	0.01	0.26
c_{dz}	0.01	0.62	0.25	0.62	0.01	0.64
c_{mz}	0.01	0.96	0.25	0.96	0.01	0.98
c_{hs}	0.01	0.21	0.25	0.21	0.01	0.22
χ^2		41.85		41.85		41.85
df		8		8		8

Table 3: Familial correlations in IQ from Bouchard & McGue (1981).

Relationship	Symbol	Number of Pairs	Correlation
MZ Twins (Together)	MZT	4672	0.86
MZ Apart	MZA	65	0.72
DZ Twins (Together)	DZT	5546	0.60
Sibs (Together)	SST	26473	0.47
Sibs Apart	SSA	203	0.24
Adopted/Biological Sibs	SFB	345	0.29
Adopted/Adopted Sibs	SFF	369	0.34
Half Sibs	HSS	200	0.31
Offspring-Parent	OPT	8433	0.42
Offspring-Midparent	OMT	992	0.50
Midoffspring-Midparent	MMT	410	0.72
Offspring-Parent Apart	OPA	814	0.22
Foster Offspring-Parent	FOP	1397	0.19
Foster Offspring-Midparent	FOM	758	0.24
Cousins	CZ1	1176	0.15
Spouses	m	3817	0.33

Table 4: The best fitting parameter sets for the data reviewed by Bouchard & McGue (1981) on familial aggregation of IQ. The estimates of the parameters in the three models are shown in roman type and the derived values are shown in *italic* type. Estimates of values not applicable for a particular model are indicated by a bar. The environmental variable e in the DP Model corresponds to the specification in equation (12)

	IP Model	DP Model	IS Model
h	0.57	0.54	0.65
b	0.52	–	0.52
β	0.47	–	0.35
$b\beta$	*0.25*	0.14	*0.18*
m	0.32	0.33	*0.32*
e	*0.58*	*0.72*	*0.53*
w	*0.11*	–	*0.04*
γ	*0.63*	*0.67*	*0.67*
r_{AA}	*0.13*	–	0.01
r_{AB}	*0.12*	–	*0.08*
r_{BB}	*0.11*	–	1.00
c_s	0.24	0.22	0.37
c_{hs}	0.22	0.17	0.09
c_{dz}	0.64	0.47	0.82
c_{mz}	0.98	0.73	1.00
χ^2	41.85	48.75	35.80
df	8	9	7
h^2	0.33	0.29	0.42

Table 5: Heritability estimates with certain subsets of the data removed. "No twins" data means that r_{MZT}, r_{MZA} and r_{DZT} are removed from data. "No raised-apart" data means that r_{MZA}, r_{SSA} and r_{OPA} are removed from data.

	Data		
	All	No twins	No raised-apart
IP Model	0.33	0.32	0.31
DP Model	0.29	0.27	0.27
IS Model	0.42	0.23	0.35

Table 6: Analysis of IQ data from Bouchard & McGue (1981) with removal of one correlation at a time. The analysis on the reduced data set repeats that used in Table 4. The number of observed data points is reduced from 16 to 15 so the overall degrees of freedom are 7–8 for the IP Model, 8–9 for the DP Model, and 6–7 for the IS Model, where the higher value obtains when a c is no longer needed (e. g. c_{dz} is not needed when r_{DZ} is removed).

	IP Model		DP Model		IS Model	
Missing Correlation	h^2	SSD	h^2	SSD	h^2	SSD
None	0.33	41.85	0.29	48.75	0.42	35.80
MZ Twins (Together)	0.33	41.85	0.29	48.75	0.44	35.71
MZ Apart	0.31	22.92	0.27	28.70	0.31	24.23
DZ Twins (Together)	0.33	41.85	0.29	48.75	0.42	35.73
Sibs (Together)	0.34	40.31	0.31	44.31	0.43	34.09
Sibs Apart	0.33	41.46	0.29	48.26	0.42	35.71
Adopted/Biological Sibs	0.33	41.48	0.30	48.17	0.42	35.11
Adopted/Adopted Sibs	0.34	36.53	0.31	38.76	0.43	30.78
Half Sibs	0.33	41.86	0.29	48.75	0.42	35.74
Offspring-Parent	0.34	39.15	0.30	47.40	0.45	32.38
Offspring-Midparent	0.33	40.32	0.30	47.58	0.43	34.16
Midoffspring-Midparent	0.32	24.75	0.29	30.72	0.40	19.27
Offspring-Parent Apart	0.34	40.96	0.31	47.13	0.42	35.63
Foster Offspring-Parent	0.33	41.79	0.29	48.47	0.42	36.07
Foster Offspring-Midparent	0.33	41.85	0.29	48.70	0.42	35.75
Cousins	0.33	41.57	0.29	48.17	0.42	35.46
Spouses	0.43	32.64	0.33	45.19	0.47	31.40

Table 7: Correlations between characters in the biological family and the foster family of adopted individuals, from Goldberger (1978).

	Observed Correlation	N	Study
Biological parent education with foster parent education	0.31	124	Leahy
	0.25	94	Leahy
	0.29	836	Leahy
	0.27	100	Skodak
Biological mother education with foster mid-parent IQ	0.20	89	Leahy
Biological mother IQ with foster mid-parent education	0.24	100	Leahy

Table 8: Correlations between the described variables in the two foster families of monozygotic twins raised apart, from Bouchard et al. (1990).

	Observed Correlation
Foster fathers' education	0.134
Foster mothers' education	0.412
Foster fathers' SES	0.267
Material possessions	0.402

Table 9: Data analysis (with r_{MMT} removed) assuming that monozygotic twins raised apart experience correlated environments. The table is organized as was Table 4.

	IP Model	DP Model	IS Model
h	0.54	0.51	0.57
b	0.50	–	0.51
β	0.51	–	0.36
βb	*0.26*	0.15	*0.18*
m	0.33	0.33	*0.33*
e	*0.62*	*0.74*	*0.59*
w	*0.12*	–	*0.11*
γ	*0.60*	*0.64*	*0.63*
r_{AA}	*0.12*	–	0.05
r_{AB}	*0.11*	–	*0.19*
r_{BB}	*0.11*	–	0.78
c_s	0.24	0.25	0.33
c_{hs}	0.20	0.18	0.08
c_{dz}	0.57	0.48	0.70
c_{mza}	0.98	0.66	1.00
c_{mz}	0.91	0.75	1.00
χ^2	4.91	9.84	5.25
df	6	7	5
h^2	0.30	0.26	0.33

Table 10: Heritability estimates and SSD values for the models assuming no common-environmental effects (with r_{MMT} removed).

	Heritability (h^2)	SSD	d. f.
IP Model	0.74	633	11
DP Model	0.82	685	12
IS Model	0.70	499	10

Table 11: Heritability estimates for the data of Bouchard & McGue (1981) as analysed in Table 9. The first three rows show our heritability estimates obtained from the full IP, DP, and IS models. The subsequent two rows provide estimates based on the parent-offspring correlation: the conditional heritability estimate, $h_c^2 = 2b_{OP|\text{culture}}$ (described in equation 88 using the parameter estimates from the IP model) and the value of Fisher's heritability, h_F^2 (see Cavalli-Sforza and Bodmer, 1971, p.546). The last three rows give measures of heritability based on twin data, Holzinger's H_H, Nichols' H_N (see Hay, 1985, p.221), and Jensen's H_J (Jensen 1989).

	IQ Data
h_{IP}^2 h_{DP}^2 h_{IS}^2	0.30 0.26 0.33
h_c^2	0.52
$h_F^2 = \frac{2r_{OPT}}{(1+m)}$	0.63
$H_H = \frac{(r_{MZT}-r_{DZT})}{(1-r_{DZT})}$	0.65
$H_N = \frac{2(r_{MZT}-r_{DZT})}{r_{MZT}}$	0.60
$H_J = r_{MZA}$	0.72

Table 12: Familial correlations in three synthetic personality characters from Eaves et al. (1989).

Relationship	Number	Correlation		
		Psychoticism	Extraversion	Neuroticism
MZ Twins	297	0.438	0.507	0.459
DZ Twins	231	0.257	0.160	0.081
Spouses	155	0.273	0.036	0.063
Parent-Offspring	533	0.125	0.207	0.117
Siblings	409	0.160	0.246	0.044
Grandparent-Offspring	45	−0.016	0.043	0.199
Aunt/Uncle-nephew	302	0.016	0.127	0.042
Cousins	104	0.008	0.093	0.052
Second Cousins	23	−0.214	0.178	−0.293
Foster Parent-Offspring	208	0.001	−0.022	0.068
Foster Siblings	52	0.016	−0.115	0.232

Table 13: Reanalysis of psychoticism data. For each model the left column records the estimates obtained without the parameters accounting for common environment and the right column records the estimates for the full model. The organization of the table is as in Table 4.

	IP Model		DP Model		IS Model	
h	0.59	0.40	0.59	0.40	0.62	0.42
b	0.00	0.71	–	–	0.54	0.61
β	0.00	0.03	–	–	0.00	0.03
βb	*0.00*	*0.02*	0.00	0.01	*0.00*	*0.02*
m	0.21	0.27	0.21	0.27	*0.27*	*0.27*
e	*0.81*	*0.59*	*0.81*	*0.91*	*0.57*	*0.67*
w	*0.00*	*0.00*	*0.59*	*0.40*	*0.00*	*0.00*
r_{AA}	*0.07*	*0.04*	–	–	0.00	0.01
r_{AB}	*0.00*	*0.08*	–	–	*0.00*	*0.09*
r_{BB}	*0.00*	*0.14*	–	–	0.93	0.60
c_s	–	0.20	–	0.08	–	0.14
c_{dz}	–	0.50	–	0.20	–	0.37
c_{mz}	–	0.81	–	0.33	–	0.58
χ^2	13.83	1.73	13.83	1.77	8.93	1.69
df	7	4	8	5	6	3
h^2	0.35	0.16	0.35	0.16	0.38	0.17

Table 14: Reanalysis of extraversion data. The organization of the table is as in Tables 13 and 4.

	IP Model		DP Model		IS Model	
h	0.69	0.63	0.69	0.63	0.69	0.60
b	0.00	0.02	–	–	0.01	0.00
β	0.01	0.14	–	–	0.08	0.16
βb	*0.00*	*0.00*	0.00	0.00	*0.00*	*0.00*
m	0.00	0.04	0.00	0.04	*0.00*	*0.05*
e	*0.72*	*0.77*	*0.73*	*0.77*	*0.72*	*0.80*
w	*0.00*	*0.00*	*0.69*	*0.63*	*0.00*	*0.01*
r_{AA}	*0.00*	*0.02*	–	–	0.00	0.14
r_{AB}	*0.00*	*0.00*	–	–	*0.00*	*0.03*
r_{BB}	*0.00*	*0.00*	–	–	0.10	0.01
c_s	–	0.05	–	0.05	–	0.04
c_{dz}	–	0.00	–	0.00	–	0.00
c_{mz}	–	0.18	–	0.18	–	0.23
χ^2	4.46	2.73	4.46	2.72	4.46	2.54
df	7	4	8	5	6	3
h^2	0.47	0.40	0.47	0.40	0.47	0.36

Table 15: Reanalysis of neuroticism data. The organization of the table is as in Tables 13 and 4.

	IP Model		DP Model		IS Model	
h	0.57	0.06	0.57	0.21	0.57	0.00
b	0.00	0.48	–	–	0.40	0.46
β	0.00	0.41	–	–	0.00	0.37
βb	*0.00*	*0.20*	0.00	0.08	*0.00*	*0.17*
m	0.01	0.07	0.01	0.06	*0.06*	*0.07*
e	*0.82*	*0.87*	*0.82*	*0.97*	*0.71*	*0.89*
w	*0.00*	*0.00*	*0.57*	*0.23*	*0.00*	*0.23*
r_{AA}	*0.00*	*0.00*	–	–	0.00	0.48
r_{AB}	*0.00*	*0.00*	–	–	*0.00*	*0.40*
r_{BB}	*0.00*	*0.02*	–	–	0.35	0.33
c_s	–	0.00	–	0.03	–	0.00
c_{dz}	–	0.00	–	0.04	–	0.01
c_{mz}	–	0.49	–	0.42	–	0.49
χ^2	23.26	5.17	23.26	5.79	22.67	4.97
df	7	4	8	5	6	3
h^2	0.33	0.00	0.33	0.04	0.33	0.00

Table 16: The parameter estimates of Eaves et al. (1989) obtained using a model similar to the DP model. The parameters β_D, b, m, Δ and w are equivalent to the parameters b, e, μ, d, and ρ, respectively, used by Eaves et al. (1989). The parameter Δ is a measure of the degree of dominance. For psychoticism and extraversion, the left columns show the model with the fewest parameters that still fits and the right columns show the parameter estimates for the full model. For neuroticism, only the full model shows a reasonable fit. The organization of the table is as in Table 4.

	Psychoticism		Extraversion		Neuroticism
h	0.60	0.47	0.68	0.67	0.00
β_D	-	-0.00	-	-0.03	0.11
b	0.80	0.74	0.73	0.71	0.76
$b\beta_D$	-	*−0.00*	-	*−0.02*	*0.08*
Δ	-	0.49	-	0.26	0.65
m	0.21	0.27	-	0.04	0.07
w	-	*−0.00*	-	*−0.02*	*0.00*
χ^2	15.7	6.4	5.2	3.6	10.8
df	9	7	10	7	7

Table 17: Heritability estimates for the data of Eaves et al. (1989). The first two estimates show the estimates of Eaves et al. (1989). The next three rows show our heritability estimates obtained from the full IP, DP, and IS models from Tables 13-15. The subsequent two rows provide estimates based on the parent-offspring correlation: the conditional heritability estimate, $h_c^2 = 2b_{OP|\text{culture}}$ (described in equation 88 using the parameter estimates from the full IP model) and the value of Fisher's heritability, h_F^2 (see Cavalli-Sforza and Bodmer, 1971, p.546). The last two rows give measures of heritability based on twin data, Holzinger's H_H and Nichols' H_N (see Hay, 1985, p.221).

	Psychoticism	Extraversion	Neuroticism
h^2	0.22	0.45	0.00
h^2_{broad}	0.46	0.52	0.42
h^2_{IP}	0.16	0.40	0.00
h^2_{DP}	0.16	0.40	0.04
h^2_{IS}	0.17	0.36	0.00
h_c^2	0.36	0.42	0.00
$h_F^2 = \frac{2r_{OPT}}{(1+m)}$	0.20	0.40	0.22
$H_H = \frac{(r_{MZT}-r_{DZT})}{(1-r_{DZT})}$	0.24	0.41	0.41
$H_N = \frac{2(r_{MZT}-r_{DZT})}{r_{MZT}}$	0.83	1.37	1.65

Table 18: Familial correlations that had a large influence on h^2. Shown are the correlations that, when deleted from the data set, led to a change of at least 0.05 in the heritability estimate, h^2. The expected value for the missing correlation is calculated on the basis of the parameters that best fit the models using the reduced data set (the estimated parameter, h^2, computed in the absence of the given correlation is provided in parentheses). For comparison, the observed correlations from Table 12 are also supplied.

Removed Correlation	Trait	Expected Correlation			Observed Correlation
		IP Model	DP Model	IS Model	
Offspring-Parent	P	0.067 (0.11)	0.067 (0.11)	0.063 (0.12)	0.125
Foster Offspring-Parent	P	0.055 (0.11)	0.063 (0.09)	0.060 (0.13)	0.001
Aunt/Uncle-niece	P	0.053 (0.19)	0.053 (0.19)	0.058 (0.23)	0.016
Spouses	E	0.161 (0.36)	0.140 (0.37)	0.103 (0.30)	0.036
Sibs Together	N	0.279 (0.09)	0.299 (0.12)	0.280 (0.05)	0.044
Adopted/Adopted Sibs	N	0.014 (0.09)	0.009 (0.08)	0.021 (0.05)	0.232

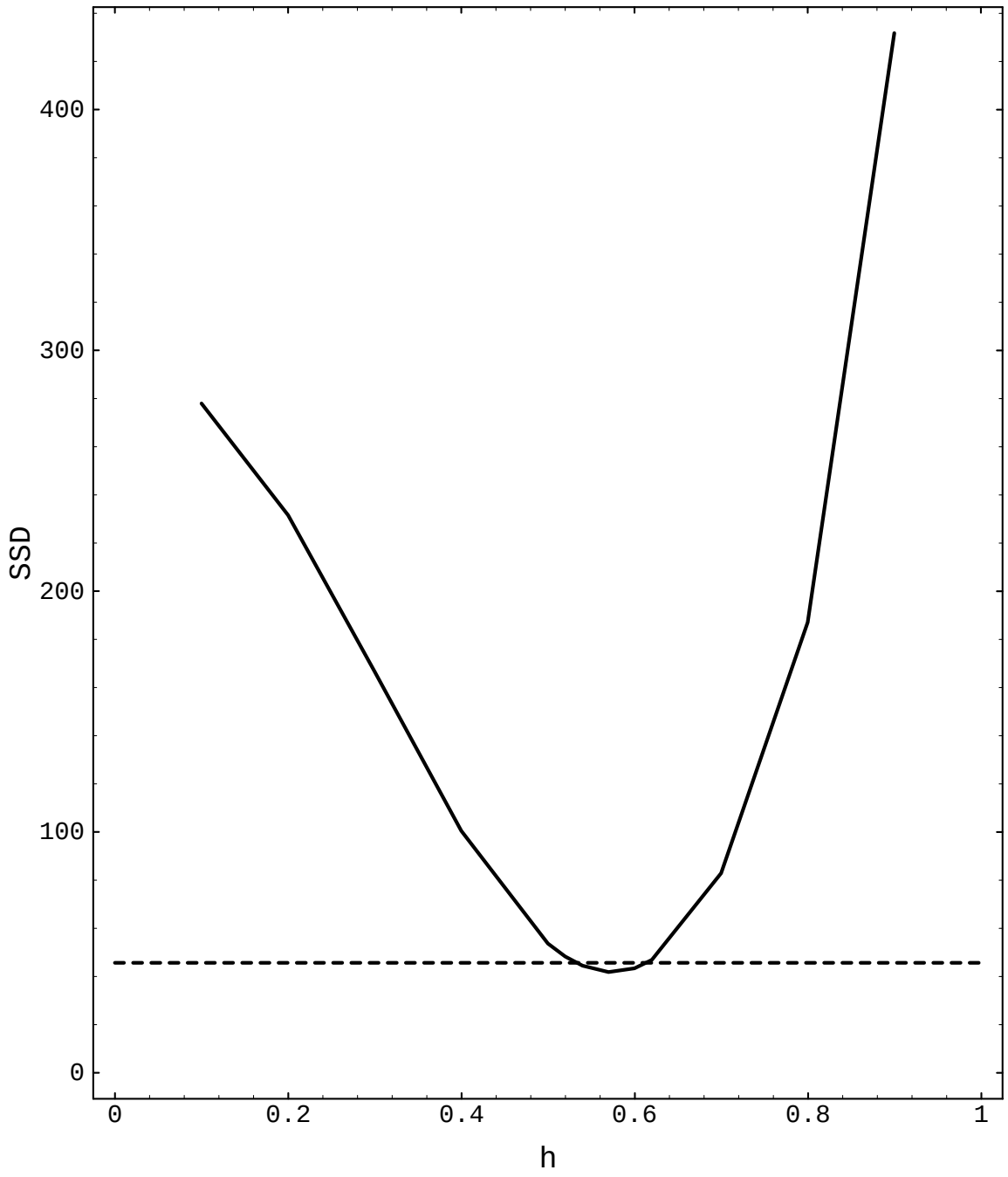


Figure 1: Dependence of the sum of squared deviations (SSD) on h in the IP Model.

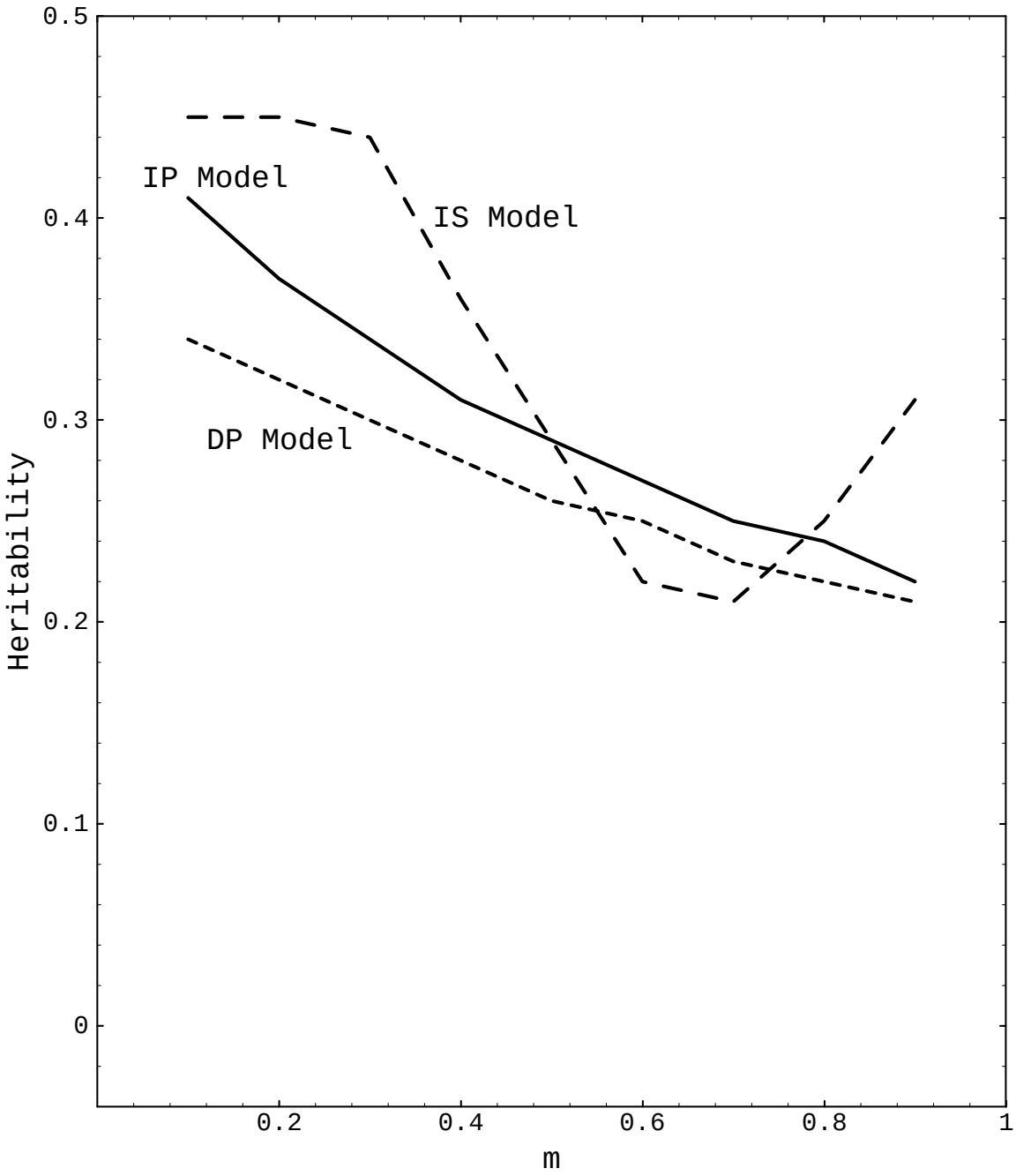


Figure 2: Heritability (h^2) as a function of m in the three models.

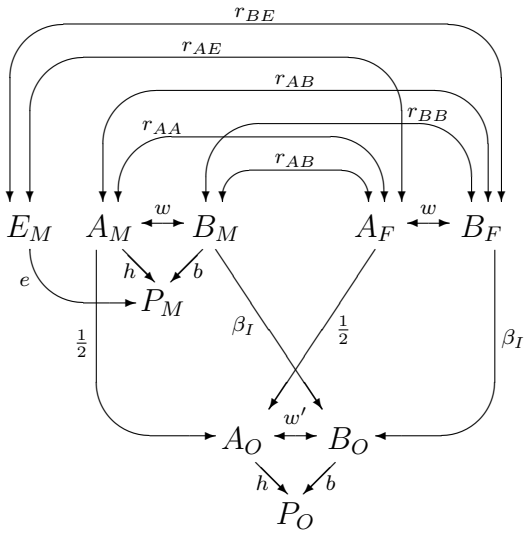


Figure 3: Path diagram describing the relation between mother and offspring in the indirect inheritance model.

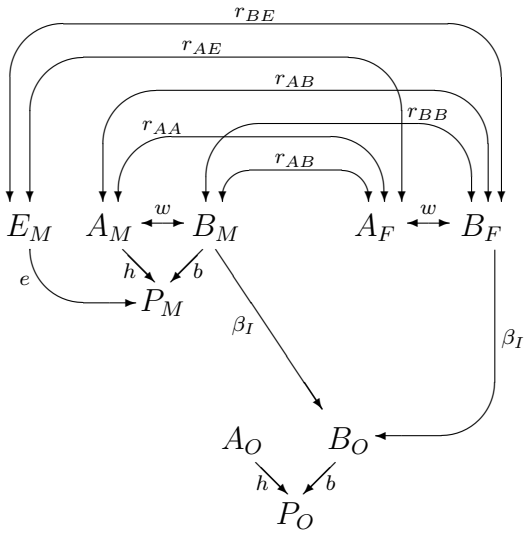


Figure 4: Path diagram illustrating the relation between a foster mother and her offspring adopted at birth in the indirect inheritance model.

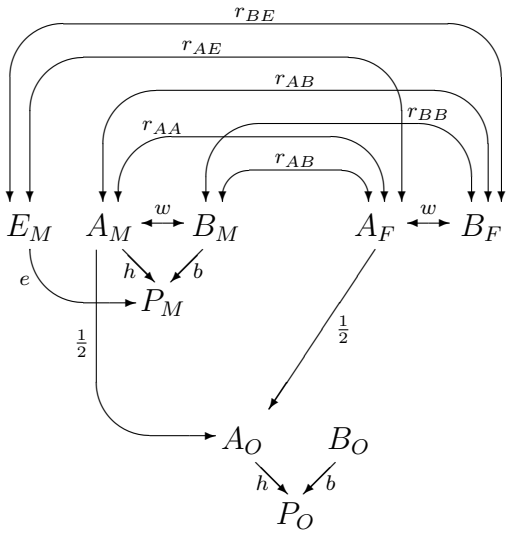


Figure 5: Path diagram illustrating the relation between a mother and her biological offspring who was adopted away at birth in the indirect inheritance model.

REFERENCES

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., and Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*, *250*, 223–228.

Bouchard, T. J. and McGue, M. (1981). Familial studies of intelligence. *Science*, *212*, 1055–1059.

Cavalli-Sforza, L. L. and Bodmer, W. F. (1971). *The Genetics of Human Populations*. W. H. Freeman and Co., San Francisco.

Cavalli-Sforza, L. L. and Feldman, M. W. (1973). Cultural versus biological inheritance: Phenotypic transmission from parents to children (A theory of the effect of parental phenotypes on children's phenotypes). *Am. J. Hum. Genet.*, *25*, 618–637.

Cavalli-Sforza, L. L. and Feldman, M. W. (1978). Dynamics and statistics of traits under the influence of cultural transmission. In N. E. Morton and C. S. Chung (Eds.), *Genetic Epidemiology* (pp. 133–143). Academic Press, New York.

Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton Univ. Press, Princeton, NJ.

Cavalli-Sforza, L. L., Feldman, M. W., Chen, K. H., and Dornbusch, S. M. (1982). Theory and observation in cultural transmission. *Science*, *218*, 19–27.

Cloninger, C. R., Rice, J., and Reich, T. (1979a). Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *Am. J. Hum. Genet.*, *31*, 176–198.

Cloninger, C. R., Rice, J., and Reich, T. (1979b). Multifactorial inheritance with cultural transmission and assortative mating. III. Family structure and the analysis of separation experiments. *Am. J. Hum. Genet.*, *31*, 366–388.

Eaves, L. J., Eysenck, H. J., and Martin, N. G. (1989). *Genes, Culture and Personality: An Empirical Approach*. Academic Press, San Diego, CA.

Ewens, W. J. (1979). *Mathematical Population Genetics*. Springer-Verlag, New York.

Falconer, D. S. (1989). *Introduction to Quantitative Genetics* (3 ed.). Longman, New York.

Feldman, M. W. and Cavalli-Sforza, L. L. (1977). The evolution of continuous variation. II. Complex transmission and assortative mating. *Theor. Pop. Bio.*, *11*, 161–181.

Feldman, M. W. and Cavalli-Sforza, L. L. (1979). Aspects of variance and covariance analysis with cultural inheritance. *Theor. Pop. Bio.*, *15*, 276–307.

Feldman, M. W. and Lewontin, R. C. (1975). The heritability hang-up. *Science*, *190*, 1163–1168.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinburgh*, *52*, 399–433.

Fisher, R. A. (1993). *Statistical Methods, Experimental Design, and Scientific Inference*. Oxford University Press, Oxford.

Goldberger, A. S. (1978). *Models and Methods in the IQ Debate: Part I revised.* Social Systems Research Institute, University of Wisconsin.

Graybill, F. A. (1961). *An Introduction to Linear Statistical Models*, volume 1. McGraw-Hill Book Co., Inc., New York.

Hay, D. A. (1985). *Essentials of Behaviour Genetics*. Blackwell Scientific Publications, Boston.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educ. Rev.*, *39*, 1–123.

Karlin, S. (1979). Models of multifactorial inheritance: I, Multivariate formulations and basic convergence results. *Theor. Pop. Bio.*, *15*, 308–355.

Karlin, S., Cameron, E. C., and Chakraborty, R. (1983). Path analysis in genetic epidemiology: A critique. *Am. J. Hum. Genet.*, *35*, 695–732.

Kirkpatrick, M. and Lande, R. (1989). The evolution of maternal characters. *Evolution*, *43*, 485–503.

Lewontin, R. C., Rose, S., and Kamin, L. J. (1984). *Not in our Genes: Biology, Ideology, and Human Nature*. Pantheon Books, New York.

Li, C. C. (1955). *Population Genetics*. University of Chicago Press, Chicago.

Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L., and Eysenck, H. J. (1986). Transmission of social attitudes. *Proc. Natl. Acad. Sci. USA*, *83*, 4364–4368.

Morton, N. E. and Rao, D. C. (1978). Quantitative inheritance in man. *Yearbook of Physical Anthropology*, *21*, 12–41.

Plomin, R., Corley, R., DeFries, J. C., and Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. *Psychological Science*, *1*, 371–377.

Rao, D. C. and Morton, N. E. (1978). IQ as a paradigm in genetic epidemiology. In N. E. Morton and C. S. Chung (Eds.), *Genetic Epidemiology* (pp. 145–194). Academic Press, New York.

Rao, D. C., Morton, N. E., and Cloninger, C. R. (1979). Path analysis under generalized assortative mating. I. Theory. *Genet. Res. Camb.*, *33*, 175–188.

Rao, D. C., Morton, N. E., Lalouel, J. M., and Lew, R. (1982). Path analysis under generalized assortative mating. II. American IQ. *Genet. Res. Camb.*, *39*, 187–198.

Rao, D. C., Morton, N. E., and Yee, S. (1974). Analysis of family resemblance. II. A linear model for familial correlation. *Amer. J. Human Genet.*, *26*, 331–359.

Rao, D. C., Morton, N. E., and Yee, S. (1976). Resolution of cultural and biological inheritance by path analysis. *Amer. J. Human Genet.*, *28*, 228–242.

Rice, J., Cloninger, C. R., and Reich, T. (1978). Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of the unitary models. *Amer. J. Human Genet.*, *30*, 618–643.

Sokal, R. R. and Rohlf, F. J. (1981). *Biometry* (2 ed.). W. H. Freeman and Co., New York.

Wright, S. (1921). Systems of mating. I. The biometric relations between parents and offspring. *Genetics*, *6*, 111–123.

Wright, S. (1931). Statistical methods in biology. *J. Am. Stat. Assoc.*, *26*, 155–163.