Behavior Genetics, Vol. 15, No. 1, 1985

Resolving the Effects of Phenotype and Social Background on Mate Selection

A. C. Heath[1,2] and L. J. Eaves[1]

Received 9 May 1984—Final 28 Aug. 1984

Traditional studies of mate selection have not addressed the question of how a marital correlation arises. The common assumption that assortative mating is based on phenotype has not been properly tested. Social background may be a major determinant of choice of spouse. We show how the collection of data on monozygotic (MZ) and dizygotic (DZ) twin pairs and their spouses, and estimation of all possible correlations between the twin pairs and their spouses, will allow these alternative hypotheses to be tested. Power simulations show that it will be feasible to resolve the contributions of phenotype and social background to mate selection for variables such as IQ, education, and attitudes for which the marital correlation is moderately high.

KEY WORDS: mate selection; mixed homogamy; twins.

INTRODUCTION

The study of monozygotic (MZ) and dizygotic (DZ) twin pairs is the most powerful method available for detecting genetic influences on human variation (Eaves, 1970; Martin *et al.*, 1978). Recently proposed extensions of the classical twin design, to include data on the parents of twins (Lytton,

---

Research reported in this paper was supported by a postgraduate studentship to ACH from the United Kingdom Medical Research Council and by Grants GM30250 and HL28922 from the National Institutes of Health. We are grateful to Drs. N. G. Martin, C. R. Cloninger, D. C. Rao, and W. E. Nance for helpful comments on early drafts of this paper.

[1] Department of Human Genetics, Medical College of Virginia, Richmond, Virginia 23298, U.S.A.

[2] Formerly Department of Experimental Psychology, University of Oxford, Oxford, England.

0001-8244/85/0100-0015$04.50/0

1977; Eaves *et al.*, 1978), the offspring of MZ twins (Nance and Corey, 1976) or of MZ and DZ twins (Hayley and Last, 1981), and the spouses of MZ and DZ twins (Eaves, 1979), permit us to advance beyond the mere study of heritability. With such extended designs we may determine more precisely the "genetic architecture" of a trait (Dobzhansky, 1972; Mather, 1966), its mode of environmental transmission (Cavalli-Sforza and Feldman, 1973; Eaves, 1976a,b; Rao *et al.*, 1976; Cloninger *et al.*, 1979), and the mechanism of mate selection for the trait (Eaves, 1973,1979; Rao *et al.*, 1979). These issues are closely related: incorrect and untested assumptions about assortative mating, for example, will usually lead to incorrect conclusions about genetic architecture and environmental transmission.

There is strong evidence of positive assortative mating, the tendency for like to marry like, for behavioral traits such as IQ (Jensen, 1978; Johnson *et al.*, 1980), attitudes and beliefs (Eaves *et al.*, 1978; Feingold, 1984), and, to a lesser extent, certain personality traits (Cattell and Nesselrode, 1967), and for related demographic variables such as educational level (Warren, 1966), socioeconomic status (Centers, 1949), and religious affiliation (Thomas, 1951). Unfortunately investigations of mate selection have not progressed beyond the tabulation of marital correlations for a list of variables. They have not sought to identify the causes of such correlations.

It is widely assumed that marital correlations arise through positive "phenotypic" assortative mating, i.e., that individuals prefer to marry or are more likely to encounter each other because they have similar phenotypes. This is the assumption which is implicit in most genetic analyses of behavioral traits (Jinks and Fulker, 1970; Eaves, 1973; Eaves *et al.*, 1978; Loehlin, 1978; Rice *et al.*, 1980). An alternative explanation is possible. Any tendency for individuals from similar social backgrounds to marry each other will also lead to a positive marital correlation.

Living in the same neighborhood and attending the same church, school, or college will increase the probability that two people will marry. If place of residence, education, or worship is purely determined by the phenotypes of the spouses themselves, this may still be considered a case of phenotypic assortative mating. If these and similar variables are directly determined, wholly or in part, by the phenotypes of the parents of the spouses, a positive marital correlation will still be generated because of the resemblance of the two sets of parents and of each set of parents and their own offspring. Such an effect of social background on mate selection will have rather different implications for genetic analysis from those of positive phenotypic assortative mating (Heath and Eaves, 1984).

Evidence for such an effect will also constitute important progress in our understanding of the psychology of mate selection.

It is not possible to infer from an observed marital correlation whether mate selection has been determined by the observed trait, some correlated feature of the individual's phenotype (Rao *et al.*, 1979; Cloninger, 1980), or social background. Direct enquiry of the spouses themselves will not be fruitful, since there is no necessary connection between the causes of behavior and the causes to which such behavior is attributed. In this paper we show how the collection of data on pairs of MZ and DZ twins and their spouses, or pairs of siblings and their spouses, will allow the resolution of the effects of phenotype and social background on mate selection. Power simulations are reported which show the sample sizes needed to resolve these effects.

A MIXED HOMOGAMY MODEL

Only a single model of "social homogamy" has been used in genetic analysis (Rao *et al.*, 1976,1979). This model has been more aptly described as a model for assortative mating based on a correlated latent variable (Cloninger, 1980). As we have shown elsewhere (Heath and Eaves, 1984), the model implies that the marital correlation for an observed variable arises because phenotype assortative mating is based on some other feature of the phenotypes of spouses, a feature which is determined by the same genetic factors ("pleiotropy") or the same environmental factors as the observed variable. We have therefore developed a new model to represent the effects of both phenotype and social background on mate selection.

In Fig. 1 we reproduce a path diagram (Wright, 1934,1968,1978) for the resemblance of siblings and their parents and spouses under a mixed homogamy model (Heath and Eaves, 1984). We assume that assortative mating is based on a complex character such as place of residence, worship, or education, which is a function of the phenotype (e.g., IQ) of the individual and of both parental phenotypes. Symbols and parameters used in Fig. 1 are as follows: P—observed phenotype; PL—complex variable on which mate selection is based; A, C, D, and E—additive genetic and familial environmental values and dominance and random environmental deviations which determine P; MO, FA, SO, DA, WI, and HU—subscripts (underlined in Fig. 1 for greater clarity) identifying mother's, father's, son's, daughter's, son's wife's, and daughter's husband's phenotypes, etc.; f—path regression of PL on individual's phenotypic value P; s—path regression of PL on maternal or paternal phenotypic value P; μ—primary marital correlation between the complex variables (e.g., ed-

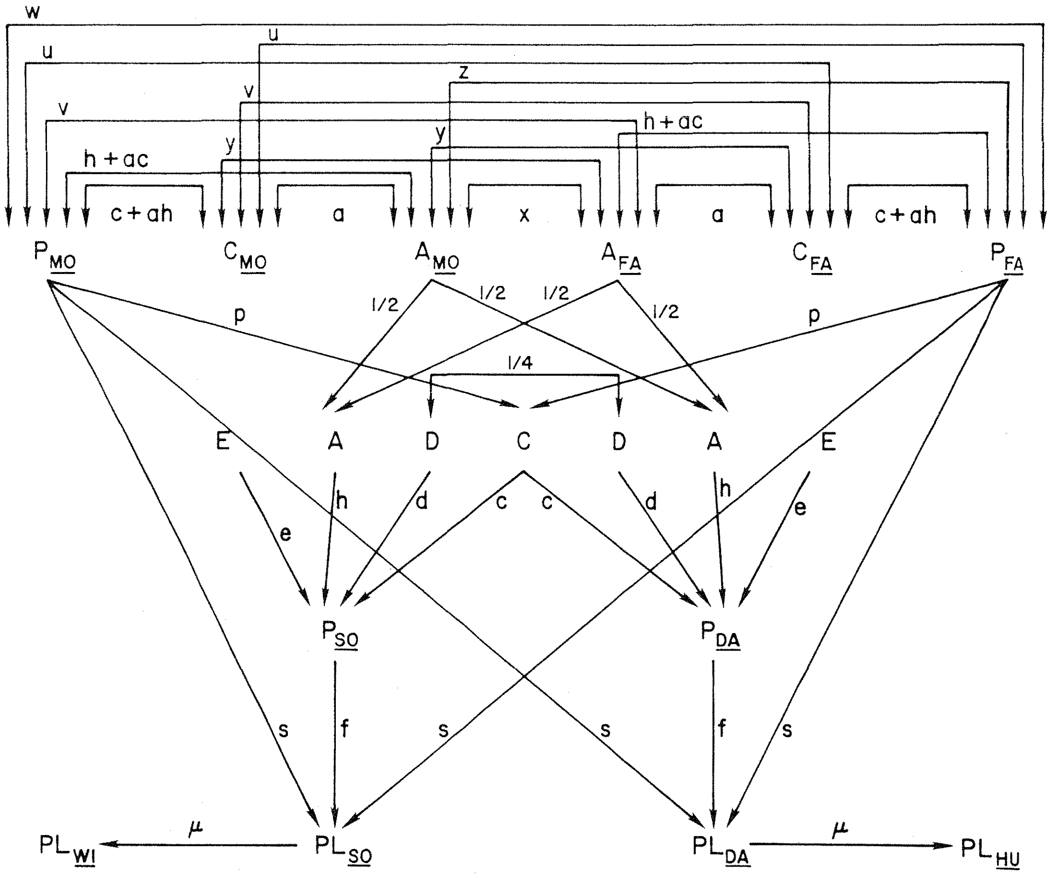

Fig. 1. Resemblance of siblings and their parents and spouses under mixed homogamy.

ucational levels) of spouses; h, c, d, and e—path regressions of P on A, C, D, and E; p—path regression of offspring familial environmental value on maternal or paternal phenotypic value; a—genotype–environmental correlation; u—correlation between phenotypic value of one spouse and familial environmental value of second spouse; v—correlation between additive genetic value of one spouse and phenotypic value of second spouse; w—correlation between phenotypic values of spouses; x—correlation between additive genetic values of spouses; y—correlation between additive genetic value of one spouse and familial environmental value of second spouse; z—correlation between familial environmental values of spouses; and ω—path regression of P on PL. In this diagram the path regressions of the complex phenotypes (PL) of the spouses on those of the twins have been represented by a single-headed arrow. This requires the use of reverse paths to derive expected correlations between other variables in one spouse and variables in the twins or cotwin's spouse (Cloninger *et al.*, 1979; Heath, 1983; Heath and Eaves, 1984). The same diagram could be drawn using copaths (Cloninger, 1980), which would obviate the need for reverse paths but still lead to the derivation of identical expected correlations. We have represented the phenotypes of the parents, as well as their additive genetic (A) and familial environmental (C) values, as ultimate variables. When both phenotype and social background contribute to mate selection, the correlation between the phenotypes of spouses cannot be expressed as a simple linear function of the correlations between their additive genetic and familial environmental values, or vice versa. Residual variables which contribute to the variance of a variable, but not to its covariance with other variables, have been omitted from the diagram. We use rectangular ($\overline{\downarrow\qquad\downarrow}$) two-headed arrows to denote correlations between ultimate variables, rather than following the traditional convention of curved two-headed arrows, since this facilitates the interpretation of complex path diagrams.

We assume that assortative mating, genetic inheritance, and cultural transmission remain constant from generation to generation, so that an equilibrium state has been reached in which the phenotypic variances and correlations between relatives do not change between generations. We may therefore, without loss of generality, assume that all measured and latent variables have been standardized to have zero mean and unit variance. This permits us to use the rules of standardized path analysis (Wright, 1934,1968,1978) to derive expected correlations between variables. Equilibrium expressions for the correlations w, v, u, x, y, z, and a will be as follows:

$$w = \mu\omega^2,$$

$$v = \mu\omega\gamma,$$

$$u = \mu\omega\epsilon,$$

$$x = v^2/w,$$

$$y = uv/w,$$

$$z = u^2/w,$$

and

$$a = p(h + ac + v),$$

where $\omega = f + hs(h + ac + v) + 2csp(1 + w)$, $\gamma = f(h + ac) + s(h + ac + v)$, and $\epsilon = f(c + ah) + 2s(1 + w)$. Since all variables have been standardized to have unit variance, the model implies inequality constraints

$$f^2 + 2s^2(1 + w) + 4fs[1/2h(h + ac + v) + pc(1 + w)] \leq 1$$

and (1)

$$2p^2(1 + w) \leq 1.$$

Table I. Expected Correlations Between Relatives[a]

Relationship	Expected correlation
MZ twins	$\rho_{MZ} = h^2 + d^2 + c^2 + 2hac$
DZ twins/full sibs	$\rho_{DZ} = \frac{1}{2}h^2(1 + x) + \frac{1}{4}d^2 + c^2 + 2hac$
MZ half-siblings	$\frac{1}{4}h^2(1 + 2\mu\gamma^2 + \mu^2\gamma^2\pi) + c^2p^2(\rho_{MZ} + 2\mu\omega\eta + \mu^2\pi\omega^2) + hcp(h + ac + \mu\gamma\eta + \mu\gamma\omega + \mu^2\gamma\omega\pi)$
First cousins	$\frac{1}{8}h^2(1 + x + 4\mu\gamma\xi + 2\mu^2\gamma^2\delta) + c^2p^2(\rho_{DZ} + 2\mu\omega\beta + \mu^2\delta\omega^2) + hcp[\frac{1}{2}h(1 + x) + ac + \mu\gamma\beta + \mu\xi\omega + \mu^2\gamma\omega\delta]$
Parent–offspring	$\frac{1}{2}h(h + ac + v) + cp\ (1 + w)$
MZ cognate uncle–nephew	$\frac{1}{2}h(h + ac + \mu\gamma\eta) + cp(\rho_{MZ} + \mu\omega\eta)$
Cognate uncle–nephew	$\frac{1}{4}h[h(1 + x) + 2ac + 2\mu\gamma\beta] + cp(\rho_{DZ} + \mu\omega\beta)$
Husband–wife	$\mu\omega^2$
MZ twin–cotwin's spouse	$\mu\omega\eta$
DZ twin–cotwin's spouse	$\mu\beta\omega$
Spouses of MZ twin pairs	$\mu^2\omega^2\pi$
Spouses of DZ twin pairs	$\mu^2\omega^2\delta$
MZ affine uncle–nephew	$\frac{1}{2}h(\mu\omega\gamma + \mu^2\pi\omega\gamma) + cp(\mu\omega\eta + \mu^2\pi\omega^2)$
Affine uncle–nephew	$\frac{1}{2}h(\mu\omega\xi + \mu^2\delta\omega\gamma) + cp(\mu\omega\beta + \mu^2\delta\omega^2)$

[a] Where $\pi = f^2(h^2 + d^2 + c^2 + 2hac) + 2s^2(1 + w) + 2fsh(h + ac + v) + 4fcsp(1 + w)$, $\eta = f(h^2 + d^2 + c^2 + 2hac) + 2s[\frac{1}{2}h(h + ac + v) + cp(1 + w)]$, $\delta = f^2[\frac{1}{2}h^2(1 + x) + \frac{1}{4}d^2 + c^2 + 2hac] + 2s^2(1 + w) + 2fsh(h + ac + v) + 4fcsp(1 + w)$, $\xi = f[\frac{1}{2}h(1 + x) + ac] + s(h + ac + v)$, $\beta = f[\frac{1}{2}h^2(1 + x) + \frac{1}{4}d^2 + c^2 + 2hac] + 2s[\frac{1}{2}h(h + ac + v) + cp(1 + w)]$, $\omega = f + hs(h + ac + v) + 2csp(1 + w)$, $\gamma = f(h + ac) + s(h + ac + v)$, and $\epsilon = f(c + ah) >+ 2s(1 + w)$.

The use of constrained optimization methods (Greig, 1980) permits us to ensure that these equality and inequality constraints are satisfied when model fitting.

EXPECTED CORRELATIONS AND PREDICTIONS

Algebraic expressions for expected correlations between MZ and DZ twin pairs and their spouses and offspring, under this mixed homogamy model, are summarized in Table I. These expectations should be multiplied by a coefficient of reliability, r, to allow for error of measurement of P. If the complex character PL is a latent variable, as will usually be the case, it will not be possible to obtain separate estimates of the parameters f, s, and μ. Instead we shall be compelled either to estimate derived parameters $f^* = f\mu^{1/2}$ and $s^* = s\mu^{1/2}$ or to assume that the variable PL is completely determined by the individual's own and his parental phenotypes, so that Eq. (1) above becomes an equality constraint. In this paper we have followed the former course.

Table II. Predictions Under Phenotypic vs. Social Homogamy[a]

Social homogamy

(i) $\rho_{ts} = \rho_{ts}' = \rho_s = \rho_s'$
(ii) $\rho_{ss} = \rho_{ss}'$
(iii) $\rho_{so} = \rho_{so}'$

Phenotypic homogamy ($r = 1$)

(iv) $\rho_{ts} = \rho_t\rho_s$
(v) $\rho_{ss} = \rho_t\rho_s^2 = \rho_{ts}\rho_s$
(vi) $\rho_{ts}' = \rho_t'\rho_s'$
(vii) $\rho_{ss}' = \rho_t'\rho_s'^2 = \rho_{ts}'\rho_s'$
(viii) $\rho_{so} = \rho_{xo}\rho_s$
(ix) $\rho_{so}' = \rho_{xo}'\rho_s'$

Phenotypic homogamy ($r < 1$)

(x) $\rho_{ts}/\rho_t = \rho_{ss}/\rho_{ts} = \rho_{ts}'/\rho_t' = \rho_{ss}'/\rho_{ts}'$

[a] ρ_t, ρ_s, ρ_{ts}, and ρ_{ss}—expected correlations between MZ twins, between husband and wife, between an MZ twin and his cotwin's spouse, and between the spouses of pairs of MZ twins; ρ_t', ρ_s', ρ_{ts}', and ρ_{ss}'—corresponding expected correlations for DZ twins or siblings; ρ_{xo} and ρ_{xo}'—expected correlation between the MZ (DZ) twin and the cotwin's offspring; ρ_{so} and ρ_{so}'—expected correlation between the spouse of an MZ (DZ) twin and the cotwin's offspring.

The utility of data on pairs or twins or siblings and their spouses for resolving the effects of phenotype and social background on mate selection can most easily be seen by deriving predictions ("scaling tests") under the two special cases where $f = 0$ (social homogamy) and $s = 0$ (phenotypic assortative mating). Scaling tests for these and other models of mate selection have been outlined elsewhere (Heath and Eaves, 1984). Predictions for the phenotypic and social homogamy models are summarized briefly in Table II and can be confirmed by simplifying the expected correlations in Table I.

If mate selection is purely determined by social background ($f = 0$), then, as noted previously (Eaves, 1979), the expected correlation between the spouses of pairs of MZ twins will be identical to the expected correlation between the spouses of pairs of DZ twins (prediction ii in Table II). Furthermore, the expected correlation between a twin and the cotwin's spouse will be the same for both MZ and DZ twin pairs and will be equal to the expected correlation between husband and wife (prediction i). This second consequence of social homogamy for the correlation between twins and their spouses had not previously been recognized. We see below that it provides the most powerful test for detecting the influence of social background on mate selection.

If mate selection is purely determined by phenotype ($s = 0$), the expected correlation between husband and wife will always be greater than that between twin and cotwin's spouse, unless the phenotypic correlation between twins is equal to unity! Under the phenotypic model with no measurement error ($r = 1$), the expected correlation between twin and cotwin's spouse will be equal to the product of the corresponding expected twin and marital correlations (predictions iv and vi), and the expected correlation between the spouses of twin pairs will be equal to the product of the square of the marital correlation and the twin correlation (predictions v and vii), for twin pairs of a given zygosity. When measurement error is important ($r < 1$), predictions iv–vii will no longer apply, but the ratio of the correlation between twin and cotwin's spouse to the correlation between twins and the ratio of the correlation between spouses of twin pairs to the correlation between twin and cotwin's spouse, should both be equal to the true correlation between spouses (i.e., corrected for unreliability of measurement; prediction x). The latter predictions will of course also be satisfied in the case where $r = 1$.

If data on the spouses of MZ and DZ twins are supplemented by data on their offspring, further tests of mechanisms of mate selection are possible. Under pure social homogamy, the expected correlation between the spouse of one twin and the offspring of the cotwin will be the same for MZ and DZ twins (prediction iii). Under the phenotypic model ($r = 1$), the expected correlation between the spouse of one twin and the cotwin's offspring should be equal to the product of the expected correlations between husband and wife and between twin and cotwin's offspring and, so, will generally be greater for MZ than for DZ kinships (predictions viii and ix). Since the absolute values of these correlations will often be quite small, these tests involving the offspring of twins will individually be much less powerful than those using twins and their spouses alone. Data on the offspring of the twins will, however, permit a more complete resolution of genetic and environmental transmission (Nance and Corey, 1976; Hayley and Last, 1981; Heath, 1983).

POWER STUDY—METHODS

From Tables I and II it is evident that data on pairs of MZ and DZ twins and their spouses will be sufficient to resolve the contributions of phenotype and social background to mate selection. It remains to be established whether this goal can be achieved with realistic sample sizes. To address this issue, we have conducted a series of power simulations, based on the work of Martin *et al.* (1978), whose original paper should be consulted for full details of methodology.

Numerical values for the expected variances and covariances of twins and their spouses were generated for a range of values of the parameters of our mixed homogamy model, including the special cases of pure social homogamy and pure phenotypic homogamy. For each twin zygosity a 4 × 4 covariance matrix was created, giving the variances and covariances of first twin, second twin, first twin's spouse, and second twin's spouse. All phenotypic variances were standardized to unity. Since we cannot estimate the cultural transmission parameter p using only data on twins and their spouses, this parameter was set to zero throughout this simulation study. In the analyses reported here measurement error was also ignored. In each case a total sample of 1000 twin pairs, each twin having one spouse, was assumed, but the proportion of twin pairs in the sample who were MZ was allowed to take values 0% (i.e., corresponding to the case where only pairs of siblings and their spouses are studied), 33.3, 50, or 66.7%.

Phenotypic and social homogamy models were fitted to the covariance matrices generated under the hypothesis of mixed homogammy, and phenotypic models were fitted to covariance matrices generated under the hypothesis of social homogamy, and vice versa. Models were fitted to the covariance matrices by maximum-likelihood covariance structure analysis (Joreskog; 1978; Eaves *et al.*, 1978; Young *et al.*, 1980). This involves minimizing the function (in standard matrix notation)

$$L = \sum N_i[\log | E_i | + tr(S\, E_i^{-1}) - \log | S_i | - p_i],$$

where S_i is the ith observed covariance matrix, E_i is the corresponding expected covariance matrix, calculated using current estimates of the free parameters of the model, N_i is the number of degrees of freedom on which the matrix is based, and p_i is the order of the covariance matrix (4 in this application). The minimum value of L obtained is distributed as chi-square, with the number of degrees of freedom equal to $[1/2\sum p_i(p_i + 1)] - q$, where q is the number of free model parameters estimated. Minimization was performed using a commercially available procedure for constrained optimization (E04UAF) using the augmented Lagrangian function method (Numerical Algorithms Group, 1978). This allowed us to ensure that the minimum obtained satisfied all equality and inequality constraints implied by the model.

For each set of parameter values under a given true model, the minimum chi-square value obtained by fitting a given false model was recorded. This chi-square statistic assesses the overall goodness of fit of the false model. A more powerful comparison of different models is possible by likelihood-ratio test (Joreskog, 1978). If L_1 is the chi-square value obtained by fitting a simple model, and L_2 is the chi-square value obtained by fitting a more complex model, then the likelihood ratio ($L_1 - L_2$) is also distributed as chi-square, with the number of degrees of freedom equal to the number of parameters of the more complex model whose values must be fixed to yield the simpler model (Joreskog, 1978). To compare alternative models by likelihood ratio test when neither one is a special case of the other, we must also fit a more complex model which includes both as special cases. If only one of the alternative models gives a significantly worse fit than the complex model, the other model will be preferred; otherwise both models will be rejected in favor of the complex model (Joreskog, 1978).

When the more complex (mixed homogamy) model is fitted to our simulated data, it will always give a perfect fit: the data were simulated taking either the more complex model or one of the alternative phenotypic or social homogamy models (which are special cases of the complex model) as the true model. The chi-square value which is obtained by fitting a false model will therefore be identical to the likelihood ratio for testing that model against the complex model. Since the likelihood-ratio statistic will be based on fewer degrees of freedom, it provides a far more powerful test of the incorrect model.

From the likelihood-ratio statistic obtained by fitting a false model to the simulated data, we may determine the number of sets of twin pairs and their spouses which would be necessary to reject the false model by likelihood-ratio test at a given significance level with a given probability, if the population parameters are those used under the true model. The minimum chi-square value obtained under the false model is used as a noncentrality parameter, λ', with the number of degrees of freedom, k, equal to the number of degrees of freedom for testing this model against the true model. The number of complete sets of twin pairs and their spouses required to reject the false model at the 5% significance level with a probability of 95% is then estimated as

$$n_p = \frac{1000\lambda}{\lambda'},$$

where λ is the noncentrality parameter $\lambda_{(0.05,0.95,k)}$ obtained from the table of noncentral chi-square (Pearson and Hartley, 1972, Table 25).

In some cases when the true value of the parameter c was zero, the estimated number of sets of twin pairs and their spouses necessary to reject a false model of assortative mating which allowed for additive gene action and familial environmental effects was greater than the number required to reject a false model which allowed for additive gene action

Table III. Number of Complete Families Required to Reject the False Social Homogamy Model When the True Model Is Phenotypic

True parameter value				% twin pairs monozygotic			
h	c	f^*	ρ^a	0	33.3	50	66.7
0.80	0.0	0.80	0.64	26	28	29	30
0.80	0.0	0.65	0.4225	79	86	90	94
0.70	0.0	0.80	0.64	21	22	22	22
0.70	0.0	0.65	0.4225	62	65	67	69
0.70	0.40	0.80	0.64	28	35	36	37
0.70	0.40	0.65	0.4225	90	111	115	119

[a] Observed marital correlation.

and purely random environmental effects. In all such cases the larger estimate of the required sample size was recorded.

RESULTS OF POWER STUDY

Table III gives the estimated number of families (i.e., complete sets of twin pairs and their spouses) required to reject a false social homogamy model when mate selection is based on phenotype, for a range of values of the parameters h and c, assuming values of the marital correlation of 0.65 and 0.4225. The former value is close to the values observed for marital correlations for educational level (Heath *et al.*, 1984) and for measures of religious and political attitudes (Feingold, 1984); the latter value is close to that which is found for individual measures of IQ. The sample size required to reject the hypothesis that mate selection is based on social background varies directly with the proportion of twins in the sample who are monozygotic, with the magnitude of the twin correlation, and inversely with the magnitude of the marital correlation. For the range of parameter values considered here, the resolution of phenotypic and social homogamy is perfectly feasible.

Table IV summarizes the total number of families required to reject a false phenotypic model when mate selection is in reality based on social background. Only a restricted range of values of the parameters h, c, and s^* are considered. Once again it is apparent that the resolution of phenotypic and social homogamy is a practically, as well as theoretically, achievable goal. Again, the necessary sample sizes vary directly with the proportion of twins who are monozygotic and with the magnitude of the twin correlations but inversely with the magnitude of the marital correlation.

Table IV. Number of Complete Families Required to Reject the False Phenotypic Model When the True Model Is Social Homogamy

True parameter value				% twin pairs monozygotic			
h	c	s^*	ρ^a	0	33.3	50	66.7
0.80	0.0	0.60	0.36	87	93	96	100
0.70	0.0	0.6332	0.18	259	277	286	296
0.70	0.40	0.6332	0.18	321	342	354	366

[a] Observed marital correlation.

Table V gives the number of sets of twin pairs and their spouses required to reject pure phenotypic and pure social homogamy models when mixed phenotypic and social homogamy is occurring. Only cases where the observed marital correlation is of the order of 0.6–0.64, or 0.4225, are considered. The resolution of mixed phenotypic and social homogamy is likely to be infeasible for smaller values of the observed marital correlation, unless available sample sizes are unusually large. We again observe that required sample sizes vary directly with the phenotypic correlation between twins and inversely with the marital correlation. Larger sample sizes are also required to reject the phenotypic model when the proportion of MZ twin pairs in the sample is high. A more complex relationship exists between the number of families required to reject the hypothesis of pure social homogamy and the proportion of twin pairs who are monozygotic: at low values of the marital correlation a high proportion of MZ twin pairs is disadvantageous, but at high values it is advantageous.

DISCUSSION

Eaves (1979) has suggested that data on the spouses of both MZ and DZ twin pairs would be necessary to resolve the contributions of phenotype and social background to mate selection. From the scaling tests given in Table II it is apparent that data on only pairs of siblings (or DZ twins) and their spouses, *or* MZ twin pairs and their spouses, would be sufficient to test the hypotheses of pure social homogamy or pure phenotypic homogamy. The results of our power study show that the resolution of phenotypic and social homogamy depends largely upon the comparison of the marital correlation and the correlation between twin and cotwin's spouse, rather than the comparison of the correlations between the spouses of MZ and those of DZ twin pairs. Under most of the conditions considered in this paper, data on pairs of siblings and their spouses will

Table V. Number of Complete Families Required to Reject the False Phenotypic and Social Homogamy Models When the True Model Is Mixed Homogamy

True parameter values					Phenotypic model, % twin pairs monozygotic				Social homogamy model, % twin pairs monozygotic			
h	c	f^*	s^*	ρ[a]	0	33.3	50	66.7	0	33.3	50	66.7
0.80	0.0	0.40	0.3434	0.64	87	92	95	98	75	73	72	71
0.70	0.0	0.40	0.4019	0.60	60	62	63	63	52	47	45	43
0.70	0.40	0.40	0.4019	0.60	70	71	72	73	115	108	105	103
0.80	0.0	0.35	0.3059	0.4225	323	362	386	413	223	233	240	248
0.70	0.0	0.35	0.3767	0.4225	201	212	219	226	160	156	154	152
0.70	0.40	0.35	0.3767	0.4225	250	264	273	281	340	341	349	358

[a] Observed marital correlation.

provide a more powerful resolution than will data on pairs of MZ and DZ twins and their spouses.

This finding is not difficult to understand. When the phenotypic correlation between twins is unity, under phenotypic as well as under social homogany, the expected correlation between a twin and his cotwin's spouse is identical to the expected correlation between husband and wife, ρ. The larger the phenotypic correlation between twins, therefore, the greater the difficulty of resolving phenotypic and social homogamy. In all the cases which we have considered in our power calculations, we have assumed that $h > 0$, so that the DZ twin correlation is less than that between MZ twins. We would therefore expect the resolution of phenotypic and social homogamy to be easier the greater the proportion of twins who are DZ in the sample. This is precisely the finding which has emerged from our power calculations.

For the behavioral geneticist who is interested not only in mechanisms of mate selection, but also in the resolution of cultural and biological inheritance, collection of data on both MZ and DZ twin pairs and their spouses will be the preferred design. Our power calculations show that with such a design, the effects of phenotype and social background on mate selection can be successfully resolved, at least for traits such as social attitudes, IQ, educational level, and socioeconomic status, with only moderately large sample sizes. The mixed homogamy model which we have considered in this paper does not of course exhaust all possible hypotheses about mate selection (see, e.g., Rao *et al.*, 1979; Eaves and Heath, 1981a,b; Heath and Eaves, 1984). To distinguish among all these different models, further extension of the classical twin design, through collection of additional data on the parents and parents-in-law of twins (Heath and Eaves, 1984), will be necessary. Additional data on the offspring of MZ and DZ twins will permit a more complete resolution of the effects of genetic and environmental transmission and, hence, a more exact determination of the genetic and cultural consequence of mate selection. However, collection of data on twin pairs and their spouses will provide the most effective starting point for an investigation of mate selection, just as the classical twin study provides the best first stage in the investigation of the familial transmission of a trait.

REFERENCES

Cattell, R. B., and Nesselroade, J. R. (1967). Likeness and completeness theories examined by 16 P.F. measures on stably and unstably married couples. *J. Pers. Soc. Psychol* **7**:351–361.

Cavalli-Sforza, L. L., and Feldman, M. (1973). Cultural versus biological inheritance: Phenotypic transmission from parents to child. *Am. J. Hum. Genet.* **25**:618–637.

Centers, R. (1949). Marital selection and occupational status. *Am. J. Sociol.* **54:**530–535.

Cloninger, C. R. (1980). Intrinsic and extrinsic structural relations by path analysis: Theory and applications to assortative mating. *Genet. Res.* **36:**135–145.

Cloninger, C. R., Rice, J., and Reich, T. (1979). 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.

Dobzhansky, T. (1972). Genetics and the diversity of behavior. *Am. Psychol.* **27:**523–530.

Eaves, L. J. (1970). *Aspects of Human Psychogenetics*, Unpublished Ph.D. thesis, University of Birmingham, Birmingham.

Eaves, L. J. (1973). Assortative mating and intelligence: an analysis of pedigree data. *Heredity* **30:**199–210.

Eaves, L. J. (1976a). A model for sibling effects in man. *Heredity* **36:**205–214.

Eaves, L. J. (1976b). The effect of cultural transmission on continuous variation. *Heredity* **37:**41–57.

Eaves, L. J. (1979). The use of twins in the analysis of assortative mating. *Heredity* **43:**399–409.

Eaves, L. J., and Heath, A. C. (1981a). On the detection of asymmetric assortative mating. *Nature* **289:**205–206.

Eaves, L. J., and Heath, A. C. (1981b). Sex-limitation and asymmetric assortative mating. In Gedda, L., Parisi, P., and Nance, W. E. (eds.), *Twin Research 3: Intelligence, Personality and Development*, Alan R. Liss, New York.

Eaves, L. J., Last, K. A., Young, P. A., and Martin, N. G. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity* **41:**249–320.

Feingold, L. F. (1984). Unpublished D.Phil. thesis, University of Oxford, Oxford.

Greig, D. M. (1980). *Optimisation*, Longman, London.

Hayley, C. S., and Last, K. (1981). The advantages of analyzing human variation using twins and twin half-sibs and cousins. *Heredity* **47:**221–236.

Heath, A. C. (1983). Unpublished D.Phil. thesis, University of Oxford, Oxford.

Heath, A. C., and Eaves, L. J. (1984). Elements of a general linear model of assortative mating (submitted for publication).

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Sundet J., Nance, W. E., Corey, L. A., and Magnus, P. (1984). No decline in assortative mating for educational level (submitted for publication).

Jensen, A. R. (1978). Genetic and behavioral effects of nonrandom mating. In Osborne, R. T., Noble, C. E., and Weyl, N. (eds.), *Human Variation: The Biopsychology of Age, Race and Sex*, Academic Press, New York.

Jinks, J. L., and Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA and classical approaches to the analysis of human behavior. *Psychol. Bull.* **73:**311–349.

Johnson, R. C., Ahern, F. M., and Cole, R. E. (1980). Secular change in degree of assortative mating for ability? *Behav. Genet.* **10:**1–8.

Joreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika* **43:**443–477.

Loehlin, J. C. (1978). Heredity-environment analyses of Jenck's IQ correlations. *Behav. Genet.* **8:**415–426.

Lytton, H. (1977). Do parents create, or respond to, differences in twins? *Dev. Psychol.* **13:**456–459.

Martin, N. G., Eaves, L. J., Kearsey, M. J., and Davies, P. (1978). The power of the classical twin study. *Heredity* **40:**97–116.

Mather, K. (1966). Variability and selection. *Proc. Roy. Soc. Ser. B* **164:**328–340.

Nance, W. E., and Corey, L. A. (1976). Genetic models for the analysis of data from the families of identical twins. *Genetics* **83:**811–826.

Numerical Algorithms Group (1978). *Fortran Library Manual, Mark 7*, NAG, Oxford.

Pearson, E. S., and Hartley, H. O. (eds.) (1972). *Biometrika Tables for Statisticians, Vol. 2*, Cambridge University Press, Cambridge.

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

Rao, D. C., Morton, N. E., Elston, R. C., and Yee, S. (1977). Causal analysis of academic performance. *Behav. Genet.* **7**:147–159.

Rao, D. C., Morton, N. E., and Cloninger, C. R. (1979). Path analysis under generalized assortative mating. I. Theory. *Genet. Res.* **33**:187–198.

Rice, J., Cloninger, C. R., and Reich, T. (1980). Analysis of behavioral traits in the presence of cultural transmission and assortative mating. *Behav. Genet.* **10**:73–92.

Thomas, J. L. (1951). The factor of religion in the selection of marriage mates. *Am. Sociol. Rev.* **16**:487–491.

Warren, B. L. (1966). A multiple variable approach to the assortative mating phenomenon. *Eugen. Q.* **13**:285–290.

Wright, S. (1934). The method of path coefficients. *Ann. Math. Stat.* **5**:161–215.

Wright, S. (1968). *Evolution and the Genetics of Populations, Vol. 1*, Chicago University Press, Chicago.

Wright, S. (1978). *Evolution and the Genetics of Populations, Vol. 2*, Chicago University Press, Chicago.

Young, P. A., Eaves, L. J., and Eysenck, H. J. (1980). Intergenerational stability and change in the causes of variation in personality. *Personal Indiv. Diff.* **1**:35–55.

Edited by C. R. Cloninger