Brief Communications

Path Analysis of Family Resemblance in the Presence of Gene-Environment Interaction

D. C. Rao\(^1\) AND N. E. Morton\(^1\)

We recently developed path analysis of familial correlations under additivity of genes and environment [1]. We recognized that nature is more complicated, but it seemed to us that testability is more important than inclusion of interactions which make a general model indeterminate.

Cavalli-Sforza and Feldman [2] have taken a different approach which reduces genetic effects to a diallelic major locus whose three genotypes have main effects (intercepts) and interactions with the familial environment (plasticity coefficients). The familial environment is assumed to be specified exactly by the midparent phenotype, which may exercise differential effects on children of different genotypes. For example, a given parental phenotype may stimulate a child of one genotype and depress another. This model has eight parameters (an intercept and plasticity coefficient for each genotype, gene frequency, and an error variance assumed the same for all genotypes). By absorbing two parameters into the population mean and variance, which are irrelevant to path analysis, the number of independent parameters is reduced to six. Random mating and absence of racial effects are assumed. Our model in this special case also has six parameters, as follows: \(h\) = the path from genotype to phenotype of child; \(c\) = the path from common environment to phenotype of child; \(ck\) = the path from common environment to phenotype of adult; \(hz\) = the path from genotype to phenotype of adult; \(r\) = the correlation between common environment and midparent genotype; and \(d\) = the path from dominance deviations to child’s phenotype, which contributes \(d^2/4\) to the sib correlation, \(d^2\) to monozygous twins, and to no other relationship. This reduces to the model of Cavalli-Sforza and Feldman for pure cultural inheritance (\(h = 0\)) or pure genetic inheritance (\(c = 0\)). If both these basic parameters are nonzero, the two models cannot be equated, and therefore our analysis cannot give an exact fit to data generated under their assumptions. It seemed to us interesting to fit their simulated data by our methods and compare the results. We assumed 1,000 pairs for

---

Received May 13, 1974.

PGL paper no. 118. This work was supported by grant GM 17173 from the U.S. National Institutes of Health.

\(^1\)Population Genetics Laboratory, University of Hawaii, Honolulu, Hawaii 96822.

© 1974 by the American Society of Human Genetics. All rights reserved.
each relationship, an enormous amount of information, especially for identical twins reared apart.

The simulated data consist of four examples with pure genetic inheritance (B1–4), four examples with pure cultural inheritance (C1–4), and 19 cases with gene-environment covariance and/or interaction. Their table 2 presented an obvious misprint (\(r_{ME} \) should be .02 for C2) and five outliers. Omitting the outliers, the remaining values were fitted well (\(\chi^2 = 62.17, \text{df} = 108\)). Including the outliers increased \(\chi^2\) by 38.01. Although we recognize that the tabulated values may be correct and our analysis unable to fit the five extreme cases, or that there may be some other misprints which we did not identify, it seems heuristic to indicate the questionable values (table 1) and perform the analysis with them included and excluded. Even after their exclusion, the sample sizes and diversity of relationships are much greater than the published material for any human trait. Omitting the five outliers, every case can be fitted well by fewer than six parameters. Thus gene-environment interaction has little chance of being recognized by goodness-of-fit tests, even in an unrealistically large body of data.

In two cases the values of \(k\) are significantly negative, which makes no sense under an additive model; both examples have negative plasticity, which means that parents of high genotype exert a depressing effect on their children, an unlikely model for most traits. In another instance the correlation of gene and environment is negative, which is surprising but not impossible. The parameter \(k\) is usually not equal to unity, and in two cases \(z\) deviates significantly. Thus our model translates an environmental effect of the parental phenotype into an intergenerational difference in the effect of heredity and common environment. Unfortunately, the deviation of \(z\) and \(k\) from unity is no test for gene-environment interaction, since it could well result from vocational, class, marital, or age effects on adults.

As advocates of path analysis, we are pleased that data generated under other assumptions are so consistent with our model. However, a critic might consider this goodness of fit a lack of specificity which invalidates attempts to infer complex patterns of genetic and cultural inheritance. To choose between these two points of view, we must try to clarify the objectives of such studies. We do not presume to anticipate the uses to which the plastic diallelic hypothesis of Cavalli-Sforza and Feldman may be put. It is unlikely that a given trait obeys their contrived model, and one is not tempted to estimate parameters of an imaginary major locus. To this caveat they might well rejoin that additivity in our model is an oversimplification, which is also unlikely to describe any trait exactly. It seems to us that the experience of biometrical genetics favors our approach over a pure major-locus hypothesis, but this argument is far from proof. What, then, do we think can be accomplished by path analysis?

First, the null hypotheses of pure cultural inheritance \(h = 0\) or pure genetic inheritance \(c = 0\) can be tested. In such a large sample as we have assumed, these tests never fail to detect a substantial value of \(h\) or \(c\) as significant. In four cases \(c = 0\) and in four cases \(h = 0\). Among the remainder, BCPV4 and BCPV11 are interesting examples in which the heritability index of Cavalli-Sforza and Feld-
### Table 1

Examples of Cavalli-Sforza and Feldman [2] \((m = 0)\)

| Model | Model of Lowest Rank Consistent with Data | Goodness of Fit | Standard Errors |
|-------|------------------------------------------|-----------------|-----------------|
|       | \(h\) | \(d\) | \(c\) | \(r\) | \(k\) | \(z\) | \(\chi^2\) | \(df\) | \(\sigma_h\) | \(\sigma_d\) | \(\sigma_c\) | \(\sigma_r\) | \(\sigma_k\) | \(\sigma_z\) |
| B1    | 1.000 | 0    | 0    | 0    | 1    | 1    | 0    | 6    | .014     | ...     | ...     | ...     | ...     | ...     | ...     |
| B2    | .577  | 0    | 0    | 0    | 1    | 1    | 0    | 6    | .013     | ...     | ...     | ...     | ...     | ...     | ...     |
| B3    | .580  | 0    | .580 | 0    | 1    | 1    | 0    | 5    | .024     | .113    | ...     | ...     | ...     | ...     | ...     |
| B4    | .535  | .535 | 0    | 0    | 1    | 1    | 0    | 5    | .039     | .059    | ...     | ...     | ...     | ...     | ...     |
| C1    | 0     | 0    | 0    | .141 | 0    | 1    | 0    | 7    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| C2    | 0     | 0    | 0    | 0    | 1    | 1    | 0    | 7    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| C3    | .566  | 0    | .566 | 0    | 1    | 1    | 0    | 5    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| C4    | 0     | 0    | 0    | 1.250| 0    | 1    | 0    | 5    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| PV1   | .506  | 0    | .506 | 0    | 1    | 1    | 0    | 5    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| PV2   | .494  | 0    | .500 | 0    | 1    | 1    | 0    | 5    | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| PV3   | .267  | 0    | .312 | 0    | 2.063| 1    | 0    | 8.43  | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BC1   | .604  | 0    | .201 | 0    | .2349| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BC2   | .616  | 0    | .345 | 0    | 1.615| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BC3   | .605  | 0    | .566 | 0    | 1.381| 1    | 0    | 2     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BC4   | .511  | 0    | .513 | .489 | 1.573| -.104| 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BC5   | .563  | .564 | .534 | 0    | 1.276| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV1 | .664  | 0    | .479 | .190 | 1.423| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV2 | .679  | 0    | .548 | .348 | 1    | .567 | 0    | 1.464| ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV3 | .711  | 0    | .388 | .414 | .379 | 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV4 | .781  | 0    | .375 | .180 | .947 | .746 | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV5 | .759  | 0    | .507 | 0    | .224 | 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV6 | .561  | 0    | .357 | 0    | .154 | 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV7 | .525  | 0    | .570 | 0    | 1.238| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV8 | .882  | 0    | .466 | -.158| 1.372| 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV9 | .725  | 0    | .395 | .553 | 1.187| .624 | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV10| .543  | .441 | .550 | .442 | .802 | 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV11| .749  | 0    | .389 | .327 | .512 | 1    | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |
| BCPV12| .795  | 0    | .367 | .207 | .873 | .857 | 0    | 1     | ...      | ...     | ...     | ...     | ...     | ...     | ...     |

Note.—\(r_{XZ}, r_{AA}, r_{S}, r_{PP}\) and \(r_{FP}\) denote the correlation coefficients between monozygous twins reared together, monozygous twins reared apart, full sibs reared together, parent and own child (reared by the parent), and parent and own child reared apart, respectively. See text for explanation of other symbols.

* \(r_{SS}\) omitted.

** \(r_{AA}\) omitted.

† \(r_{PP}\) omitted.
man (based on the variance among genotypes reared by their own parents) is null, yet identical twins reared apart are slightly correlated. Fitting $h$ together with the parameters shown in table 1, we obtain $h = .125 \pm .079$ for BCPV4 and $.114 \pm .084$ for BCPV11, which are positive, as they should be, but not significant. In all other cases we do not fail to reject a false null hypothesis. Case BCPV7, which also has a negative correlation between intercept and plasticity but a greater correlation between identical twins reared apart, has a significant heritability index by both models. Of course, a smaller sample would be less decisive, and then a false null hypothesis might be tenable (not proven). This, however, is a general statistical problem, not peculiar to family resemblance. Thus our analysis seems satisfactory for testing a null hypothesis about $c$ or $h$ in the presence of gene-environment interaction.

A second objective is to indicate the strengths of determination by cultural and genetic inheritance. Various indices of heritability and common environment are given in table 2 and plotted by the first two eigenvectors (after centroid adjustment)

| TABLE 2 |
| --- |
| INDICES OF HERITABILITY AND COMMON ENVIRONMENT |
| | Heritability | Common Environment |
| 1. Holzinger $(r_{MZ} - r_{SS})/(1 - r_{SS})$ | 1. $r_{MZ} - r_{AA}$ |
| 2. $2(r_{MZ} - r_{SS})$ | 2. $1 - (r_{PF}/r_{DP})$ |
| 3. $H^*$ | 3. $1 - H - \sigma^2/V_T$ |
| 4. $r_{AA}$ | 4. $c^2$ |
| 5. $2r_{PF}$ | 5. $c^2 + 2grc$ |
| 6. $h^2$ | 6. $1 - r_{AA} - \sigma^2/V_T$ |
| 7. $h^2 + d^2$ | |
| 8. $h^2 + d^2 + 2grc$ | |

NOTE.—$H$, $\sigma^2$, and $V_T$ denote, respectively, the heritability, random error variance, and total phenotypic variance under the model of Cavalli-Sforza and Feldman [2]. See text for explanation of other symbols.

* From Cavalli-Sforza and Feldman [2].

in figures 1 and 2. Here the outliers are included, but a similar topology is obtained when they are excluded. For heritability (fig. 1), two ad hoc indices comparing monozygous twins and siblings, including the one suggested by Holzinger [3], have low correlations with the others. The indices appropriate under our model, additive heritability ($h^2$) and broad heritability ($h^2 + d^2$), never appreciably exceed the index $H$ of Cavalli-Sforza and Feldman, which does not separate genotypic effects from the environment of children reared by their own parents and can be calculated only with prior knowledge of the genetic mechanism. The best estimate of heritability under their model is given by $r_{AA}$, the correlation of identical twins when genes and environment are randomized. This index, which does not include gene-environment interaction, is highly correlated with $h^2 + d^2 + 2grc$ and is never substantially less than $h^2$. Cavalli-Sforza and Feldman did not define an index of familial environment, but $C = 1 - H - \sigma^2/V$ would be consistent with their
A third objective is to detect gene-environment interaction and covariance. While the covariance is explicitly included in path analysis, interaction is so confounded with random environment that it must often fail to be detected by path analysis and presumably by other methods. As long as the first two objectives are satisfied, it is of little concern to us that interactions are poorly resolved.

Layzer [4], ignoring the treatment of gene-environment covariance which began with Wright [5], has emphasized complexities of gene-environment interaction. He asserts (1) that gene-environment covariance invalidates heritability analysis.
(false) and (2) that a linear model cannot exactly represent an interactive system (true). Even if interaction is greater in nature than in the model of Cavalli-Sforza and Feldman, there is no reason to suppose that biological or cultural inheritance will be exaggerated by path analysis which includes them both.

The data were fitted by the computer program SUPERVAR, a generalization of our initial program (COMVAR [1]), to include dominance, extramarital children, and families of any size [6]. As emphasized by Cavalli-Sforza and Feldman and earlier authors, adopted children are valuable for separating cultural and genetic inheritance. Increasing restrictions on human "experimentation" make studies of adopted children less feasible than in the past, but half-sibs acknowledged or detected through genetic tests provide the same kind of information.

SUMMARY

This analysis suggests that an additive model suffices to reveal the main features of complex inheritance, both cultural and biological, even in the presence of gene-environment interaction. Gene-environment covariance is included in the analysis and presents no difficulty. Neither heritability nor common environment is exaggerated by the model.

REFERENCES

1. RAO DC, MORTON NE, YEE S: Analysis of family resemblance. II. A linear model for familial correlation. Am J Hum Genet 26:331-359, 1974
2. CAVALLI-SFORZA LL, FELDMAN MW: 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, 1973
3. HOLZINGER SJ: The statistical evaluation of nature and nurture. J Am Stat Assoc 30:274-280, 1935
4. LAYZER D: Heritability analyses of IQ scores: science or numerology? Science 183:1259-1266, 1974
5. WRIGHT S: Statistical methods in biology. J Am Stat Assoc 26:155-163, 1931
6. RAO DC, MACLEAN CJ, MORTON NE, YEE S: Analysis of family resemblance. V. Height and weight in northeastern Brazil. In preparation

Employment Opportunities

POSITION OPEN

Assistant professor, statistical human genetics, University of Washington Center for Inherited Diseases. Ph.D. with strong independent research experience in statistical aspects of human genetics, computer experience, publications, some teaching. Send resume to Dr. A. G. Motulsky, Chairman of Search Committee, Division of Medical Genetics, University of Washington, Seattle, Washington 98195. An equal opportunity–affirmative action employer.