Br. J. math. statist. Psychol. (1977), 30, 185–197 *Printed in Great Britain*

An application of the analysis of covariance structures to the psychogenetical study of impulsiveness

L. J. Eaves, N. G. Martin and Sybil B. G. Eysenck

A method based on Jöreskog's approach to the analysis of covariance structures is applied to the genotype–environment analysis of the covariation of four aspects of impulsiveness in male and female monozygotic and dizygotic twins. The data are consistent with a simple model which assumes additive gene action, random mating and environmental effects within families. Further, it is shown that genotypic factor loadings can be regarded as constant multiples of corresponding environmental loadings. Trait-specific sex interactions are detected which suggest that some mechanism of sex-limitation could contribute to specific variation.

1. Introduction

The majority of multivariate behavioural studies seek a model for the phenotypic relationships between variables without reference to the possible underlying distinction between genetical and environmental determinants of individual differences. In this paper we demonstrate how the analysis of covariance structures (Jöreskog, 1973) can be exploited to combine in the same analysis a model for the psychological relationships with a causal genotype–environmental model. The approach, which is completely general, can be illustrated effectively with reference to four aspects of impulsiveness which are consistent with a simple psychogenetical model.

2. The data

Eysenck & Eysenck (1977) have suggested that impulsiveness is capable of resolution into four correlated primary factors which have been termed 'impulsiveness in the narrow sense' (Imp_n), 'non-planning', 'risk taking' and 'liveliness'. Questionnaire measurements of these factors are shown to correlate differently with questionnaire measurements of psychoticism, extraversion and neuroticism.

Copies of an experimental questionnaire were mailed to adult twins on the Institute of Psychiatry Twin Register (Kasriel & Eaves, 1976). Completed questionnaires were received from 588 pairs of twins for whom the distribution of zygosity and sex is summarized in Table 1. From the 52 impulsiveness items in the questionnaire 40 were

Table 1. Composition of twin sample employed in the analysis of impulsiveness

	Number of pairs	
	Monozygotic (MZ)	Dizygotic (DZ)
Male pairs (M)	144	52
Female pairs (F)	233	83
Unlike sex pairs (OS)	—	75

selected which best represented the four component factors of impulsiveness. Three illustrative items from each scale are given in Table 2. The numbers of items contributing to scores on each factor were: Imp_n (12); risk-taking (10); non-planning (12); liveliness (6). An angular transformation was applied to the raw scores for each factor to secure additivity but the improvement was not marked, probably on account of the relatively small number of items contributing to each of the component scales.

Table 2. Illustrative items for four impulsiveness factors

Items	Keyed response (Yes/No)
Factor: Impulsiveness (narrow sense)	
Do you often do things on the spur of the moment?	Y
Do you often get involved in things you later wish you could get out of?	Y
Do you usually think carefully before doing anything?	N
Factor: Risk	
Would you prefer a job involving change, travel and variety even though it might be insecure?	Y
Would you enjoy parachute jumping?	Y
Would you enjoy fast driving?	Y
Factor: Non-planning	
Do you think an evening out is more successful if it is unplanned or arranged at the last moment?	Y
Would you make quite sure you had another job before giving up your old one?	N
When you go on a trip do you like to plan routes and timetables carefully?	N
Factor: Liveliness	
Do you usually make up your mind quickly?	Y
Are you usually carefree?	Y
Are you slow and unhurried in the way you move?	N

3. Data summary

The mean squares and mean products between and within pairs were computed for the four variables for each of the four like-sex twin groups separately. For the male–female pairs the mean vector corresponding to the overall sex difference was also extracted from the intrapair variation. For each of the five twin groups the linear regression on age was partialled out of the variation and covariation between pairs. The age correction thus excludes from the variation between pairs any general linear trend with age but does not extract any interaction between age differences and genetical or environmental differences. Eaves & Eysenck (1976) have shown that such interactions may be detected in principle if they have a systematic component, otherwise they will remain confounded with the main effects of genetical and environmental influences but will not constitute a major problem for the interpretation of individual differences.

The corrected mean squares and mean products are given in Table 3. The d.f. take account of the corrections made for age and sex. Recognizing the constraints imposed by the symmetry of the ten matrices, the basic data summary comprises 100 statistics (ten matrices, each with ten free statistics). By considering the psychological theory implicit in the choice of the variables and the causal theory implicit in the design of the

experiment we hope to account for these in terms of a model which will lead to a significant reduction of the data.

Table 3. Age-corrected mean product matrices within and between twin pairs for four impulsiveness scales†

	Between-pairs mean product matrices						Within-pairs mean product matrices					
Twin type	d.f.		IMPN 1	RISK 2	NONP 3	LIVE 4	d.f.		IMPN 1	RISK 2	NONP 3	LIVE 4
MZF	231	1	0·12041	0·05049	0·03926	0·04673	233	1	0·05344	0·01945	0·01381	0·02299
		2	0·05049	0·12487	0·04456	0·03785		2	0·01945	0·05960	0·01312	0·01472
		3	0·03926	0·04456	0·06827	0·03286		3	0·01381	0·01312	0·03072	0·01545
		4	0·04673	0·03785	0·03286	0·17454		4	0·02299	0·01472	0·01545	0·09528
MZM	142	1	0·08964	0·03427	0·01683	0·02782	144	1	0·07445	0·03766	0·02829	0·02863
		2	0·03427	0·10136	0·02636	0·02756		2	0·03766	0·09791	0·02760	0·01528
		3	0·01683	0·02636	0·05672	0·01235		3	0·02829	0·02760	0·05021	0·02387
		4	0·02782	0·02756	0·01235	0·12959		4	0·02863	0·01528	0·02387	0·11747
DZF	81	1	0·10729	0·04840	0·03176	0·03082	83	1	0·04787	0·01216	0·01178	0·02489
		2	0·04840	0·08159	0·02951	0·01655		2	0·01216	0·03764	0·00877	0·02283
		3	0·03176	0·02951	0·06455	0·02149		3	0·01178	0·00877	0·03307	0·00579
		4	0·03082	0·01655	0·02149	0·19970		4	0·02489	0·02283	0·00579	0·08473
DZM	50	1	0·09077	0·04154	0·02206	0·01277	52	1	0·07478	0·01300	0·01298	0·00996
		2	0·04154	0·07101	0·02986	0·02515		2	0·01300	0·06791	0·01651	0·00583
		3	0·02206	0·02986	0·04510	0·03040		3	0·01298	0·01651	0·02939	0·01024
		4	0·01277	0·02515	0·03040	0·13465		4	0·00996	0·00583	0·01024	0·08983
DZOS	73	1	0·07418	0·02009	0·01796	0·01044	74	1	0·08180	0·03712	0·02634	0·02307
		2	0·02009	0·07473	0·02242	0·02658		2	0·03712	0·07394	0·01977	0·02212
		3	0·01796	0·02242	0·05869	0·01858		3	0·02634	0·01977	0·04465	0·00214
		4	0·01044	0·02658	0·01858	0·10020		4	0·02307	0·02212	0·00214	0·12258

† Key to labels for primary factors: IMPN = impulsiveness in the 'narrow sense'; NONP = non-planning; RISK = risk-taking; LIVE = liveliness.

4. The model

In a recent analysis of the causes of variation in impulsiveness and sociability, Eaves & Eysenck (1975) showed that the pattern of individual differences found in male and female monozygotic (MZ) twins and in like-sex and unlike-sex male and female dizygotic (DZ) twin pairs was consistent with a simple model which assumed that the variation and covariation of the two traits were due primarily to the additive effects of many genes and the effects of environmental influences which were largely specific to individuals rather than common to families. Furthermore, the consistency over sexes of particular estimates and the ability of the model to encompass data on unlike-sex pairs without additional parameters implied that the causes of variation in extraversion and its components did not depend substantially on sex. We suppose that the phenotypic variation for the four traits may be related to a fairly simple model invoking a single factor common to the four variables (impulsiveness in the broad sense) and components specific to each of the variables. However, by including twins in the design of our study the opportunity arises to go beyond such a simple treatment of the structure of

phenotypic variation into an analysis of its causal basis. Thus, combining the simple causal model for impulsiveness advanced by Eaves & Eysenck (1975) with the simple factorial model proposed by Eysenck & Eysenck (1977), it is possible to discover whether the factorial unity of impulsiveness applies with equal force to the genetical and environmental determinants of the trait.

The very simple model of Eaves & Eysenck will be specified in which it is implied that individual differences in impulsiveness are consistent with an explanation in terms of additive gene action, random mating and environmental effects which are unique to individuals rather than shared with other members of the same family. A fuller discussion of the assumptions of such a model can be found in the references (Martin & Eysenck, 1976; Eaves *et al.* 1977). Many of the available data on personality are consistent with such a model.

For a single variable, we may write our expectation for the total phenotypic variance in terms of our simple model thus:

$$\sigma_p^2 = \tfrac{1}{2}D_R + E_1.$$

D_R has been defined by Mather & Jinks (1971), and reflects the additive contribution of genetical differences to variation. E_1 represents the contribution of environmental differences specific to individuals (i.e. the environmental variance within families). This, it is emphasized, is the simplest model for the joint action of genetical and environmental factors. The model can be complicated in a variety of ways to include, for example, the contribution of non-additive genetical effects, assortative mating, common environmental effects, genotype–environmental interaction and genotype–environmental covariance, but the undue elaboration of the model should be avoided until the simpler explanations have been rigorously discounted.

In more general terms, for multiple variables under the same system of causation we may write for the phenotypic covariance matrix

$$\Sigma_p = \tfrac{1}{2}(\Delta\Delta' + \mathrm{D}^2) + \mathrm{HH}' + \mathrm{E}^2.$$

Of course, we do not have the phenotypic covariance matrices. Instead we have the matrices of mean products within and between pairs for the different types of twins. The contributions of the genetical and environmental factors to the different matrices differ and can be obtained from genetical theory. It is the fact that the contributions of the various causal factors differ between matrices that enables us to attempt a resolution of genetical and environmental components of phenotypic covariation. For twins we have the following expectations:

$$\begin{aligned}
\Sigma_{B\mathrm{MZ}} &= \Delta\Delta' + \mathrm{D}^2 + \mathrm{HH}' + \mathrm{E}^2, \\
\Sigma_{W\mathrm{MZ}} &= \mathrm{HH}' + \mathrm{E}^2, \\
\Sigma_{B\mathrm{DZ}} &= \tfrac{3}{4}(\Delta\Delta' + \mathrm{D}^2) + \mathrm{HH}' + \mathrm{E}^2, \\
\Sigma_{W\mathrm{DZ}} &= \tfrac{1}{4}(\Delta\Delta' + \mathrm{D}^2) + \mathrm{HH}' + \mathrm{E}^2.
\end{aligned}$$

The subscript B denotes a matrix of mean products *between* pairs, W denotes a matrix of *within-pair* mean products.

These are a simple extension of the expectations for the mean squares for a single variable given elsewhere (e.g. Eaves & Eysenck, 1977). We have made no distinction between male and female twins, or between like-sex and unlike-sex dizygotic twins in writing the initial expectations since our null hypothesis embodies the assumption that the genetical and environmental components of variation do not depend on sex.

For the present case there are four variables, and our factor model anticipates that there will be only one common factor. The vectors Δ and H will, therefore, be four-

element column vectors of genetical and environmental loadings, and D^2 and E^2 will be 4×4 diagonal matrices containing the corresponding specific variances. This suggests that our first model might attempt to explain the 100 raw statistics by reference to 16 parameters. A further simplification is proposed, however, which amounts to specifying the assumption that the common factor structure for the genetical and environmental covariation is identical such that a disturbance in the common factor has the same relative effects on all the variables, whether the origin of the disturbance be genotypic or environmental. Initially, therefore, the factor loadings are constrained so that the genetical loadings (Δ) are all a simple scalar function of the environmental loadings (H) thus:

$$\Delta = b\mathrm{H}.$$

The scalar, b, bears a simple relationship to the heritability of the common factor, given that both the factorial and the causal aspects of our model are appropriate. The model, therefore, has 13 parameters in all: four factor loadings; one scalar; four genetical specifics and four environmental specifics. In principle, further reductions might be possible by imposing constraints on the relative values of specific components, but this does not seem appropriate in the absence of any theoretical justification.

Although we have, by the nature of the particular problem, considered only the simplest experimental design and specified only the most basic genotype–environmental factor model, the approach is general in that virtually any model which can be written for a single variable can be cast in a form applicable to multiple variables, so that hypotheses can be tested and parameters estimated if an adequate experimental design is ensured.

5. Testing the model

In several papers Jöreskog (e.g. 1973) has developed and applied the conceptual, statistical and numerical methodology for problems very similar to ours. We have used an approach very similar to his, adapting it somewhat to the needs of our particular class of problem. The approach is described fully by Martin & Eaves (1977). Generally, our data will consist of k matrices of mean products. We may write S_i for the ith matrix, having N_i d.f. Given some model for the S_i we may compute the expected values Σ_i, being positive definite, for particular values of the parameters of the model. When the observations are multivariate normal we may write the log likelihood of obtaining the k observed independent S_i as:

$$\log L = -\tfrac{1}{2} \sum_{i=1}^{i=k} N_i[\log|\Sigma_i| + \mathrm{tr}\,(\mathrm{S}_i \Sigma_i^{-1})]$$

(omitting the constant term).

For a given model we require the parameter estimates which maximize $\log L$. Given maximum-likelihood estimates of our parameters we may test the hypothesis that a less restricting model (i.e. one involving more parameters) does not significantly improve the fit by computing

$$\chi^2 = 2(L_0 - L_1), \tag{1}$$

where L_1 is the log-likelihood obtained under the restricted hypothesis (H_1) and L_0 is the log-likelihood obtained under the less demanding hypothesis (H_0). The H_0 we shall adopt in practice is that which assumes that as many parameters are required to explain the

data as there are independent mean squares and mean products in the first place, i.e. $\hat{\Sigma}_i = S_i$ for every i. In this case we have simply:

$$L_0 = -\tfrac{1}{2}\sum_{i=1}^{i=k} N_i[\log|S_i| + p].$$

When we have k matrices the χ^2 has $\frac{1}{2}kp(p+1) - m$ d.f. where m denotes the number of parameters estimated under H_1 and p is the number of variables.

6. Numerical method

The likelihood is maximized by attempting to minimize $-\log L$ for a given model. There are many numerical methods for doing this. A variety of these methods has been implemented by the Numerical Algorithms Group (1974) and we employed the most flexible of their FORTRAN routines, E04HAF, for constrained minimization. The routine has the advantage of allowing the user certain flexibility in the choice of method. In particular, minimization can be based on evaluation of the function values alone (in this case the values of $-\log L$ and any functions used in specifying constraints on parameter values), or minimization can be assisted by computation of first derivatives or of first and second derivatives of the function. Furthermore, differentiation can proceed numerically or can be programmed precisely by the user. For our problem the Powell 64 method was used which relies only on the evaluation of the functions themselves, since coding the first derivatives was tedious and their approximate numerical evaluation consumed excessive computer time. A constrained minimization routine was used because of the need to ensure the $\hat{\Sigma}_i$ are all positive definite. For our simple example these constraints should be automatically satisfied, providing we estimate D and E rather than D^2 and E^2. The problem thus reduces to an unconstrained problem in our case. However, in problems which are factorially more complex, further constraints may be required to ensure that the correlation matrix of the factor loadings is positive definite. This may be done numerically in several ways, e.g. by constraining the eigenvalues or the leading minor determinants to be positive. The E04HAF routine uses a penalty function technique due to Lootsma (1972) for constraining the estimates in the required region.

In order that the estimates of a satisfactory model might be interpreted more rigorously, their covariance matrix is required. This is the inverse of the matrix of the second derivatives of the log-likelihood with respect to the maximum-likelihood parameter estimates. Jöreskog (1973) has given second derivatives of the log-likelihood for related problems involving models for single covariance matrices. We followed his approach, constructing first the matrix of second derivatives with respect to all the parameters, fixed and free, then striking out the rows and columns corresponding to fixed parameters, and finally combining the information on those parameters which are constrained to be equal. Martin & Eaves (1977) describe the method for obtaining the covariances of the estimates in further detail.

7. Results

Estimates obtained from fitting the simple model which assumed a single genetical factor with loadings proportional to those of a single common environmental factor are given in Table 4. These loadings, it should be remembered, are scaled to reproduce the covariance matrices. The log-likelihood on this hypothesis was 3867·10 for 13 parameters. If we were to estimate the parameters of a model involving the maximum of 100 parameters (a perfect fit solution), we would obtain a value of 3919·37 for the log-likelihood.

Equation (1) thus yields $\chi_{87}^2 = 104{\cdot}54$. Although this value is barely significant ($P = 0{\cdot}097$), the fit is relatively poor. Standard errors of the estimates are not cited at this stage because we are not satisfied with this model.

Table 4. Maximum-likelihood parameter estimates for model assuming identity of factor structure for genetical and environmental covariation

	Genetical†		Environmental	
Trait	Loading	Specific‡	Loading	Specific‡
IMPN	0·162	0·181	0·142	0·186
RISK	0·161	0·202	0·141	0·150
NONP	0·120	0·149	0·105	0·135
LIVE	0·126	0·278	0·111	0·249

† Genetical loadings are a constant multiple (1·1387) of the corresponding environmental loadings. The loadings are scaled to reproduce the phenotypic *covariance* matrix, not the correlation matrix.
‡ Specific *standard deviations* (not variances) are given to facilitate comparison with loadings.

We turn, therefore, to modifications of the basic explanation which could improve the fit of the model. There are several possibilities. We could seek additional common factors. This would seem unwise with only four variables. We could seek explanations which involve effects other than D_R and E_1, such as more complex environmental effects or more subtle genetical effects. Such explanations would be less parsimonious and would be inconsistent with what has already been found for the impulsiveness scale of the PEN (Eaves & Eysenck, 1975). One simple possibility would be to relax the constraint that the genetical factor loadings should be a constant multiple of the environmental loadings. This would amount to saying that the phenotypic correlation between the components of impulsiveness does not represent adequately the joint action of genetical and environmental effects and that there is not just one underlying factor but two different factors, one of which is substantially genetically determined, and the other of which reflects the structure of environmental influences. On fitting such a model the log-likelihood was 3870·40 giving $\chi_{84}^2 = 97{\cdot}94$ ($P = 0{\cdot}142$). The three additional parameters lead to a reduction in chi-square of 6·60 for 3 d.f. ($P = 0{\cdot}086$) suggesting a slight but not very significant improvement in fit.

The slight, but not very striking, evidence of some differences between genetical and environmental factor structure led us to revert to the previous model for the factor structure and we started to examine the specific variation. By a process of tentative model-fitting to the data on sexes separately, but leaving out the unlike-sex pairs, we obtained an indication that, although the factor loadings seemed fairly consistent over sexes, the values obtained for the specific variances, especially the specific genetical variances, differed quite markedly between males and females. This suggested that the genetical determinants of trait-specific variation were different in the two sexes. If this were the case we would expect the common factors to contribute to the covariation of male–female pairs, but we would expect the specific genetical variances to take different values in males and females and to make no contribution to the covariance of unlike-sex twins. Thus, a final model was fitted which differed from the initial model in only the following features. Specific genetical variances were fitted which depended on sex, with the further specification that these were genetically quite distinct in the two sexes.

This amounts to saying that the genetical component of the trait-specific variation can be best approximated by a model which assumes quite different genes are expressed in males from those expressed in females. The model is thus that described above except that we have slightly different expectations for the opposite-sex (OS) pairs, as follows:

$$\Sigma_{BOS} = \tfrac{3}{4}\Delta\Delta' + \tfrac{1}{4}(\mathbf{D}_m^2 + \mathbf{D}_f^2) + \mathbf{H}\mathbf{H}' + \mathbf{E}^2,$$
$$\Sigma_{IFOS} = \tfrac{1}{4}\Delta\Delta' + \tfrac{1}{4}(\mathbf{D}_m^2 + \mathbf{D}_f^2) + \mathbf{H}\mathbf{H}' + \mathbf{E}^2.$$

$\mathbf{D}_m^2$ and $\mathbf{D}_f^2$ denote the specific additive genetical variances for males and females respectively. In the expectation for like-sex pairs (above) we merely substitute $\mathbf{D}_m^2$ for $\mathbf{D}^2$ in the males and $\mathbf{D}_f^2$ for $\mathbf{D}^2$ in the expectation for female twins.

The model now has four factor loadings, with one constant relating genetical and environmental factors, four specific environmental components and eight specific genetical components (four for each sex), making 17 parameters in all. The log-likelihood was now 3875·12, giving an overall $\chi_{83}^2 = 88{\cdot}50$ ($P = 0{\cdot}319$) indicating a good fit and a marked improvement ($\chi_4^2 = 16{\cdot}04$, $P = 0{\cdot}003$) on the original model which assumed consistency over sexes of the specific genetical variation.

The estimates of the parameters of this, our final, model are given in Table 5. Because we have every reason to be satisfied with the fit of our model we have computed the standard errors of the estimates as a guide to their significance. These were obtained inverting the matrix of second derivatives of the log-likelihood with respect to the maximum-likelihood estimates of the parameters. Martin & Eaves (1977) describe how the covariance matrix of the estimates is obtained in such cases.

Table 5. Maximum-likelihood estimates for parameters of the final model allowing for the apparent effects of sex-limitation on specific genetical variation

	Factor loadings		Specific standard deviations		
			Genetical		
Trait	Genetical	Environmental	Female	Male	Environmental
IMPN	0·161	0·142	0·188	0·180	0·181
	(0·012)†	(0·010)	(0·021)	(0·030)	(0·009)
RISK	0·158	0·140	0·189	0·000	0·200
	(0·012)	(0·010)	(0·022)	(28·658)	(0·009)
NONP	0·118	0·105	0·143	0·123	0·149
	(0·009)	(0·008)	(0·017)	(0·025)	(0·007)
LIVE	0·127	0·113	0·245	0·284	0·274
	(0·013)	(0·011)	(0·027)	(0·037)	(0·011)

† Estimated standard errors given in parentheses.

We can see that the estimate of the genetical specific variance for the risk-taking factor is zero for males. This accounts for the very large standard error of the estimate and suggests either that there is truly no specific genetical variance for this aspect of impulsivity in males or that in this one instance an alternative explanation of individual differences is warranted. However, since the overall fit of the multivariate genotype–environmental model is good we would be loath to explain away such a specific anomaly which could have arisen by chance. Had we wished we could have repeated the estimation procedure having set this particular parameter to zero from the start (see Martin &

Eaves, 1977), but this is unlikely to lead to any substantive change to our conclusions in this example.

8. Discussion

We have shown that the data are consistent with the view that the covariance structure of impulsiveness is due to a single underlying factor which is affected jointly by genetical and environmental effects. By showing that the genetical and environmental loadings are proportional to one another we have, in effect, showed that the ratio of variation due to common genetical factors to that due to the common environmental factor is consistent over all variables. Thus, there is a common factor, which we may call impulsiveness in the broad sense, whose heritability is a simple function of the ratio, b, of the genetical and environmental loadings, given that the model fits. The fact that b differs significantly from zero ($b = 1{\cdot}1296 \pm 0{\cdot}1601$) indicates that the genetical loadings are jointly significant and justifies our attempts to estimate the proportion of the common factor variance which is due to genetic factors.

If we write δ for any one of the four genetical factor loadings, and η for the corresponding environmental loading we have $\delta = b\eta$ or $\delta^2 = b^2\eta^2$. Since the data give us no reason to doubt the adequacy of our simple genotype–environmental model we may estimate the 'heritability' of the common factor from

$$h^2 = \tfrac{1}{2}\delta^2/(\tfrac{1}{2}\delta^2 + \eta^2),$$

which is

$$h^2 = \tfrac{1}{2}b^2/(1 + \tfrac{1}{2}b^2).$$

Substituting for b gives the estimate of the proportion of the *common factor variance* which is genetically determined as 0·3895. This value does not require any correction for unreliability since we presume that sampling error will contribute only to the specific components of variation in the four measurements.

Similarly, we may use our parameter estimates to assess the relative contribution of genetical and environmental differences to the specific variation of the four measurements. By fitting various models, we have established that the sexes differ in the genetical mechanism responsible for specific variation, so we are compelled to give separate estimates of the 'specific heritabilities' for each sex.

Writing d_i^2 for a typical genetical specific and e_i^2 for the corresponding environmental specific we have

$$h_i^2 = \tfrac{1}{2}d_i^2/(\tfrac{1}{2}d_i^2 + e_i^2).$$

These estimates are in Table 6. In fact, with the exception of the risk factor for which males show no specific genetical variation in contrast to the females who show significant genetical specific variation, the values are fairly consistent over sexes. If our only concern were to simplify the numbers we could now try fitting to both sexes the same numerical values to certain of the specific variances. However, we have not done so because our model-fitting has suggested a quite distinct causal basis for the specific variation in two sexes. Any attempt to constrain the specifics to be equal over sexes would, therefore, be unduly arbitrary.

We notice that the values for the heritability are comparable in magnitude for that obtained for the common factor. However, we should remember that errors of measurement, whilst they should not contribute to environmental variation in the common factor (given independent errors), would be expected to contribute to the

specific environmental variation. It is clearly desirable to obtain some estimate of the measurement error, in order to assess the relative contributions of unreliability and 'real' environmental effects to the environmental variation within families. This can be approximated since the analysis is based on scores which have been subjected to the angular transformation for which a theoretical error variance is available. For a scale

Table 6. Proportion of variation specific to four impulsiveness scales attributable to genetical factors

Trait	Females	Males
IMPN	0·350	0·380
RISK	0·329	0·000
NONP	0·332	0·332
LIVE	0·144	0·146

consisting of n items of equal difficulty and given local independence the theoretical error variance takes the value $1/4n$. If the items of a scale are not all equally difficult then this estimate of error will be larger than the true value.

Table 7 shows estimates of the specific environmental variation for each trait with the appropriate theoretical errors for scales of the corresponding length. The difference between the two sets of estimates is an estimate of the 'true' specific environmental variation which is due to factors other than errors of measurement. In every case, rather more than half of the measurable specific environmental variation within families seems to be attributable to errors of measurement. Indeed, in the case of the non-planning factor, we conclude that virtually all the detectable specific environmental variation is due to sampling error in the scores.

Table 7. Analysis of the contribution of measurement error to specific variation

		Estimated contribution to specific environmental variance			'Heritability' of specific variance corrected for measurement error‡	
Trait	No. of items	'Error' variance†	'True' specific environmental variance	Proportion due to 'error'	Females	Males
IMPN	12	0·021	0·012	0·63	0·60	0·57
RISK	10	0·025	0·015	0·63	0·54	0·00
NONP	12	0·021	0·001	0·95	0·91	0·88
LIVE	6	0·042	0·033	0·56	0·48	0·55

† The error variance is estimated as $1/4n$ (see text).
‡ Cf. uncorrected values in Table 6.

Finally, Table 8 gives a summary of the contributions of the different sources of variation to the four measurements in each sex. The values were obtained from the

estimates in Table 5. For each sex, the phenotypic variance for each trait was calculated by substituting the appropriate parameter estimates in

$$\sigma_p^2 = \tfrac{1}{2}(\delta^2 + d^2) + \eta^2 + e^2.$$

The parameters δ, d, η and e take the values appropriate for the particular sex and variable in question. Then the contribution from each source is expressed as a proportion of the total for that variable and sex.

Table 8. Summary of the relative contributions of common and specific genetical and environmental effects to the total variation in each of the components of impulsiveness

		Proportion due to genetical effects			Proportion due to environmental effects		
Sex	Trait	Common	Specific	Total	Common	Specific†	Total
Female	IMPN	0·155	0·211	0·366	0·241	0·392	0·633
	RISK	0·139	0·218	0·357	0·199	0·444	0·643
	NONP	0·138	0·219	0·357	0·203	0·440	0·643
	LIVE	0·064	0·101	0·165	0·238	0·596	0·835
Male	IMPN	0·158	0·245	0·404	0·197	0·399	0·596
	RISK	0·173	0·000	0·173	0·272	0·554	0·827
	NONP	0·146	0·231	0·377	0·158	0·465	0·623
	LIVE	0·059	0·094	0·153	0·296	0·551	0·847

† Error variation has not been deducted from the contribution of specific environmental factors.

Many other summary statistics could be derived from the estimates in Table 8, including traditional heritability estimates for the individual variables. Adding together the contributions from the genetical and environmental common factor yields the familiar communality estimate for each variable. Adding the genetical contributions due to common and specific variance for each variable in turn, we have the usual heritability estimate applicable to each variable as it would be derived in any equivalent univariate analysis of the individual scales.

9. Conclusion

The specific application presented here should not be allowed to obscure the generality of our approach. Given adequate family groupings and sufficiently strong psychological expectations it is possible to formulate and test a psychogenetical model for individual differences which embodies both the causal and psychological components of any theory of individual differences. Although we have been content to fit a single common factor, because this was inherent in our choice of measurements, the approach can be extended to the estimation of additional correlated or uncorrelated factors as long as appropriate constraints are specified or fixed values are assigned for certain of the factor loadings (Jöreskog, 1973). In the present case, we have shown that the multivariate structure of impulsiveness is consistent with a causal explanation which assumes additivity of gene action and the absence of environmental differences between families. Further, the specific variation appears to reflect a different causal basis in the two sexes. From the

psychological viewpoint we suggest that the genetical and environmental covariation in the different measures of impulsiveness can be regarded as mediated by the same structure which is jointly influenced by genetical and environmental factors.

As well as yielding the maximum-likelihood estimates of the factor loadings and the specific standard errors, with all their consequent desirable properties, the approach enables us to test the adequacy of the fitted model and provides us with estimates of the standard errors of the parameter values, so that the margin of error attached to individual estimates can be assessed.

Even within the scope of our simple example we have seen some of the flexibility of the method because we have been able to decide between a model which allows for sex-limitation of the expression of the gene loci responsible for specific variation and one which assumes that the same specific gene effects operate in both sexes. We have been able to show that the genetical loadings can be regarded as constant multiples of the environmental loadings, and that the contribution of the common factors to trait covariation is consistent over sexes. If our simple model had failed we could have attempted to explain the failure in terms of non-additive gene action, or environmental differences between families, or assortative mating. Providing the investigator possesses the ingenuity to write the appropriate model and collect the right data, the possibilities for the causal analysis of trait covariation in quantitative genetical terms seem extensive. With data on relatives other than twins it would be possible to study still greater subtleties of the mechanism underlying the multivariate structure of individual differences.

Acknowledgements

We thank Mrs J. Kasriel for her careful organization of data collection for this project. The first two authors acknowledge the support of a programme grant for psychogenetical research from the British Medical Research Council. We thank our colleagues, including Professors H. J. Eysenck, J. L. Jinks and Mr Owen White for discussion and encouragement.

References

Eaves, L. J. & Eysenck, H. J. (1975). The nature of extraversion: A genetical analysis. *J. Person. soc. Psychol.* **32**, 102–112.

Eaves, L. J. & Eysenck, H. J. (1976). Genotype x age interaction for neuroticism. *Behav. Genet.* **6**, 359–362.

Eaves, L. J. & Eysenck, H. J. (1977). A genotype–environmental model for psychoticism. *Adv. Behav. Res. Ther.* **1** (in press).

Eaves, L. J., Last, K. A., Martin, N. G. & Jinks, J. L. (1977). A progressive approach to non-additivity and genotype–environmental covariance in the analysis of human differences. *Br. J. math. statist. Psychol.* **30**, 1–42.

Eysenck, H. J. & Eysenck, S. B. G. (1976). *Psychoticism as a Dimension of Personality.* London: Hodder & Stoughton.

Eysenck, S. B. G. & Eysenck, H. J. (1977). The place of impulsiveness in a dimensional scheme of personality. *Br. J. soc. clin. Psychol.* **16**, 57–68.

Jöreskog, K. J. (1973). Analysis of covariance structures. In P. R. Krishnaiah (ed.), *Multivariate Analysis III. Proceedings of the Third International Symposium on Multivariate Analysis.* New York: Academic Press.

Kasriel, J. & Eaves, L. J. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *J. biosoc. Sci.* **8**, 263–266.

Lootsma, F. A. (1972). A survey of methods for solving constrained minimization problems via unconstrained minimization. In F. A. Lootsma (ed.), *Numerical Methods for Non-linear Optimization.* New York: Academic Press.

Martin, N. G. & Eaves, L. J. (1977). The genetical analysis of covariance structure. *Heredity* **38**, 77–95.

Martin, N. G. & Eysenck, H. J. (1976). Genetic factors in sexual behaviour. In H. J. Eysenck (ed.), *Sex and Personality*. London: Open Books.
Mather, K. & Jinks, J. L. (1971). *Biometrical Genetics: The Study of Continuous Variation*. London: Chapman & Hall.
Numerical Algorithms Group (1974). E04HAF. In NAG Library Mark 5 Manual: Oxford University: NAG Central Office.

Received 11 September 1976; revised version received 3 May 1977

Requests for reprints should be addressed to L. J. Eaves, Department of Genetics, University of Birmingham, Birmingham B15 2TT.
N. G. Martin is at the same address.
S. B. G. Eysenck is at the Institute of Psychiatry, University of London.