Behavior Genetics, Vol. 2, No. 2/3, 1972

SHORT COMMUNICATION

The Multivariate Analysis of Certain Genotype–Environment Interactions

L. J. Eaves[1]

Received 17 March 1971—Final 8 Sept. 1971

The multivariate analysis of sums and absolute differences for scores of pairs of separated identical twins on multiple measures provides an indication of the generality of certain types of genotype–environment interaction for behavioral traits. The method is illustrated by analysis of part of the data from the Michigan Twin Study.

INTRODUCTION

The correlation between the sums and absolute differences for scores of pairs of separated identical twins has been shown to provide a test of a systematic linear relation between genotypic differences and sensitivity to environmental differences for quantitative characters in human populations (Jinks and Fulker, 1970). Consideration of isolated variables, however, does not distinguish a general genetic component of sensitivity to environmental differences from components specific to particular tests and situations. Simultaneous study of multiple variables does permit the generality of such genotype–environment interactions to be investigated. The method suggested by Jinks and Fulker may be extended to the multivariate case to permit a statistical test of the minimum number of dimensions required to represent these interactions and to allow their structure to be investigated.

This work was carried out while the author was in receipt of a postgraduate research studentship from the British Science Research Council and is part of a project in psychogenetics supported by the British Medical Research Council.

[1] Department of Genetics, University of Birmingham, Birmingham, England.

METHOD

For a sample of n pairs of separated identical twins each measured on p variates, the sums and absolute differences are calculated. A partitioned matrix of order $2p$ is formed thus:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}$$

where $\mathbf{R}_{11}$ represents the correlations among the p sets of pair sums, $\mathbf{R}_{22}$ the correlations among the p sets of differences, and $\mathbf{R}_{12} = \mathbf{R}'_{21}$ the correlations between the sums and differences. If there are no linear genotype–environment interactions, $\mathbf{R}_{12}$ will be null; if there are interactions specific to particular variables, some of the diagonals of $\mathbf{R}_{12}$ will have significant entries, and some of the off-diagonals of $\mathbf{R}_{12}$ will be significant in addition if any of the interaction is common to two or more variables.

The minimum dimensionality of the interaction may be tested by calculating the significance levels associated with the p canonical correlations of **R**. The structure of the interaction is best revealed by factor analysis of **R**. If sums and differences load on independent factors, there is no evidence that a genotype's sensitivity to the environment is linearly related to its mean performance over different environments. Factors loading simultaneously on sums and differences define traits of interaction between genotypic and environmental differences. If many tests show high loadings for sums and differences on a few factors, it may be concluded that the interaction displays a high degree of generality. Nonlinear and unsystematic interactions will not be detected in this way, but may be detected as common factors of $\mathbf{R}_{22}$ since this matrix represents the variation of intrapair differences around their mean value. This may not be necessarily related linearly to genotypic differences and still reflect genotype–environment interaction.

EXAMPLE

The method is illustrated by analysis of the scores of 19 pairs of female identical twins on a small selection of motor skills. The twins, which constituted part of the sample for the Michigan Twin Study (Vandenberg, 1962), were reared together. Scores were obtained for tests of mirror-drawing and card-sorting with left and right hands separately. Vandenberg describes the tests in more detail. Since the twins were reared together, the sum–difference correlation is biased by any interaction of common environmental influences with environmental differences within families. A preliminary genetic analysis including a sample of fraternal twins revealed that there was no need

Table I. The Partitioned Matrix for Genotype–Environment Interaction Analysis

Sums				Differences			
Mirror		Cards		Mirror		Cards	
RH	LH	RH	LH	RH	LH	RH	LH
1.00	0.29	0.14	0.20	−0.01	0.07	−0.37	−0.14
	1.00	0.22	0.09	−0.14	0.67	−0.08	−0.13
		1.00	0.88	0.08	0.35	0.68	0.58
			1.00	−0.04	0.11	0.60	0.68
				1.00	−0.10	0.19	−0.16
					1.00	0.06	−0.08
						1.00	0.72
							1.00

to invoke common environments to account for any of the observed variation for these measurements. The 8 × 8 **R** matrix is given in Table I. The subsequent analysis assumes that the distribution of the absolute intrapair differences is normal.

Even with this small sample, there are highly significant sum–difference correlations ($p<0.01$) for mirror-drawing with the left hand and card-sorting with left and right hands. Examination of Table I suggests that the two card-sorting tests form a unit for which environmental influences interact jointly with common genetic influences. The interaction for mirror-drawing is specific to the left hand and independent of that for card-sorting.

Two of the four canonical correlations (Table II) may be regarded as nonsignificant since the two smallest canonical roots are homogeneous when tested by chi-square ($\chi^2_{(2)} = 4.05$, $p>0.05$). It is therefore possible to represent the systematic genotype–environment interaction in a minimum of two dimensions.

Maximum likelihood estimates of the factor loadings cannot be obtained because the solution of the equations for this matrix required that a number of the specific variances be negative. The first two principal components were extracted instead and rotated orthogonally by varimax to improve the approximation to simple structure. The loadings are given in Table III.

Table II. Canonical Roots and Correlations

Root	Correlation	χ^2	df	P
0.76	0.87	19.55	7	<0.01
0.56	0.75	11.12	5	~ 0.05
0.20	0.44	4.05	4	>0.05
0.08	0.29	—	—	—

Table III. Rotated Components of **R**

			I	II
Sums	Mirror	RH	−0.08	0.54
		LH	0.04	0.88
	Cards	RH	0.90	0.29
		LH	0.89	0.15
Differences	Mirror	RH	0.03	−0.22
		LH	0.16	0.79
	Cards	RH	0.87	−0.24
		LH	0.84	−0.22
Proportion of variance			0.39	0.24

The first factor relates high mean performance to high variability for both card-sorting tests, and the second shows the common variation of genotypic and intrapair differences for left-hand mirror-drawing. Clearly motor skills such as the two considered here are factorially complex so it is to be expected that there is little evidence of a general factor relating the genes determining sensitivity to the environment to those responsible for mean performance over a range of environments. With so few tests, it is difficult to distinguish between interaction arising as a result of arbitrary test parameters, such as the difficulty of the tasks, and interaction between genotypic and long-term environmental differences. The data, however, do illustrate a procedure which could make such discrimination possible with a larger sample of twins and a judicious selection of tests.

ACKNOWLEDGMENTS

The author is indebted to Professor S. G. Vandenberg for supplying the data which illustrate this paper and to Professor J. L. Jinks for his continued advice and encouragement.

REFERENCES

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

Vandenberg, S. G. (1962). The hereditary abilities study: Hereditary components in a psychological test battery. *Am. J. Hum. Genet.* **14:** 220–237.