Ann. Hum. Genet., Lond. (1975), 39, 219
Printed in Great Britain

The inheritance of scholastic abilities in a sample of twins

II. Genetical analysis of examination results

By N. G. MARTIN

Department of Genetics, University of Birmingham, Birmingham

There is now strong evidence from many different studies that the greater part of variation in IQ, in western white societies at least, is genetically determined. The evidence for educational achievement is based on fewer studies and is more conflicting. In this paper we detect and estimate genetic components of performance in the different subjects of a school examination and then seek to discover whether these components are all measures of the same general ability or whether they also include variance due to independently inherited specific abilities.

THE DATA

In a previous paper (Martin & Martin, 1975) we described how 149 pairs of twins were ascertained from the records of the Intermediate Examination of the Public Examinations Board of South Australia in 1967 and 1968. No important biases in the ascertainment of the sample or diagnosis of zygosity were found and it was hoped that representative estimates of parameters describing variation in examination performance could be obtained. In all there were 98 pairs of same-sex and 39 pairs of opposite-sex twins with examination results for analysis. A detailed breakdown of the sample by year, sex and zygosity is given in the previous paper.

The main problem presented by these data is that although each candidate was required to take about seven subjects for examination, there was a choice of over 25 different subjects of which only English was compulsory. Thus, English and Mathematics were taken by quite large numbers of students while fewer took languages and very few twins were found taking technical and commercial subjects. The picture is further complicated by some different subjects being offered in the two years in which data were collected. Thus Physics and Chemistry of 1967 are blended into a combined science course examined as Science 1 and 2 in 1968. However, the syllabuses for the other subjects were substantially the same in both years.

It would be interesting to know whether genetic differences influenced the choice of subjects taken by an individual. We can sum over all the subjects offered, the number of cases in which both members of a pair took a subject (concordant) and those in which only one member took it (discordant). Then we may compare the proportion of concordant pair-subjects for the two twin types. In so doing we must remember that since most individuals do the same number of subjects, a single discordant event gives rise to two discordant pair-subjects. In comparing MZ and DZ twins therefore, we must halve the number of discordant pair-subjects for each class. Because the sex difference within opposite-sex pairs will inevitably influence the choice of subjects, we shall exclude these pairs from the comparison. Table 1 shows the number of concordant and discordant pair-subjects for each year, sex and zygosity.

There is no significant difference in concordance between MZ and DZ twins for either sex in each of the two years. This suggests that genetic differences have little influence on the choice of subjects taken, this being mainly determined by cultural or E_2 differences between pairs.

Table 1. *Concordance for pair-subjects by year, sex and zygosity*

(Number discordant is halved in calculation of contingency chi square.)

	1967				1968			
	MZ female	DZ female	MZ male	DZ male	MZ female	DZ female	MZ male	DZ male
Concordant	76	80	81	42	99	59	66	83
Discordant	31	33	4	11	33	24	13	30
χ^2_1	0·03		3·15		0·07		1·00	

Table 2. *Analysis of variance of English scores of opposite-sex pairs*

Source	D.F.	M.S.	F	
Between pairs	33	275·2233	2·35	$P < 0·01$
Between sexes	1	317·7794	2·86	$P = 0·10$
Within pairs residual	33	111·2946		

Table 3. *Concordance of tested pairs for school class during the Intermediate year*

Classes at Intermediate	MZ	DZ	
Same	21	8	29
Different	7	11	18
	28	19	47

$\chi^2_1 = 3·88,\ 0·02 < P < 0·05.$

As described in the previous paper, 47 of the same-sex pairs were personally interviewed so that we could check on zygosity diagnosis. At the same time each twin had his stature measured and did the Australian Council of Educational Research Higher Test, a group IQ test comprising a 15 min language test and a 20 min quantitative test. These two characters are also included in the genetical analysis.

The raw examination scores for each subject are scaled out of 100 with a mean of 45 points and standard deviation of approximately 15. The scaled marks are not quite normally distributed as about 5 % of students fall above the second standard deviation. This slight distortion was introduced to facilitate the awarding of prizes and scholarships. In practice, inspection of the skewness and kurtosis coefficients of the observed distributions reveals very few significant departures from normality. These would only be of concern if they introduced spurious genotype–environment interactions or genetical non-additivity which, as we shall see, they do not.

A copy of the raw data for all twin pairs has been deposited with the National Auxiliary Publications Service.

Ideally one would like to treat each twin type of each sex in each year as a separate class so that the heterogeneity in fitting a single set of parameters to statistics from all ten classes could be assessed directly. Unfortunately the numbers in each class are too small to permit this so we must pool some of the classes, first checking that their means and variances are homogeneous.

We can reduce the ten classes to five by combining the data for 1967 and 1968 twins in each class except for the science subjects which examine different syllabuses in the two years. F-tests of the total variances reveal three out of 38 comparisons significant at the 5 % level and we may regard these as due to chance. Differences in means are similarly trivial.

The numbers in each twin class are still small so we are further forced to pool male and female

MZ twins into a single MZ class and the three DZ classes into a single DZ group. There is a large significant difference in mean height between males (175·2 cm) and females (163·3 cm) so it is not possible to pool data for this character over sexes, but for all the other characters there were no substantive differences in means between sexes. The five variances (only four for IQ) of the 11 remaining characters were tested for homogeneity using Bartlett's test and one comparison (Chemistry) was significant at the 5 % level.

As a test of certain assumptions made in fitting simple biometrical models to these data, we checked that there were no significant differences between the total variances of the combined MZ and DZ classes. The means of the two groups were also compared and it was found that MZ twins scored significantly higher in English and History but it is unlikely that these substantively small differences in means will affect analysis of the variances.

Finally, we checked that in pooling the opposite-sex pairs with the other DZ classes we were not introducing a variance term due to the sex difference within the pair. This would tend to inflate the DZ within-pair variance and increase the apparent heritability. Table 2 shows an analysis of variance of English scores in the opposite-sex pairs in which sex is taken as a main effect.

It can be seen that there is no consistent effect of sex on English scores, and in the other five subjects (French, History, Geography, Mathematics 1 and 2) in which numbers were large enough to give sufficiently powerful tests the sex effect was even smaller. Thus, we are justified in pooling the three DZ classes for each of the subjects.

We can use the same analysis to test another possible within-pair effect. During the interview with the 47 same-sex pairs the twins were asked which of them was born first and every pair gave an unequivocal answer. Although this opinion is far from reliable, we looked for a consistent within-pair effect of parity but no appreciable differences between first and second born were found.

It could be argued that if DZ twins are in different school classes and have different teachers more frequently than do MZ twins then this could inflate the environmental variance within DZ pairs. The members of all but two of the 110 same-sex pairs attended the same school in their Intermediate year. The co-twins of the 47 tested pairs were asked whether they were in the same or different classes in their Intermediate year. Their answers are compared for MZ and DZ twins in Table 3.

There is just significant disproportionality which indicates that DZ twins are more often placed in different classes than MZ. This greater difference in placement probably reflects greater genetic dissimilarities although the separation may reinforce such differences. If the sample were larger it would be interesting to compare genetic analyses of data from twins in the same and in different school classes.

If possible, one should test for the presence of genotype–environment correlation and the different types of genotype–environment interaction before applying simple models of inheritance to family data. Jinks & Fulker (1970) provide a number of tests for these complications but without MZ twins reared apart we can only test for the presence of interaction between genotypes (G) or between-family environments (E_2) and within-family environments (E_1). They suggest that a relationship between pair-sums and absolute within-pair differences of MZ twins reared together is evidence of GE_1 or $\mathrm{E}_1\mathrm{E}_2$ interactions, although in the unlikely event that the interaction term is greater than the E_1 contribution then such a relationship will not be found.

Table 4. *Analysis of variance for second degree polynomial regression of absolute pair differences on pair sums for MZ English scores*

Source	D.F.	M.S.	F
Linear term	1	8·8972	0·31
Quadratic term	1	1·7306	0·06
Deviation about regression	48	28·4999	
Total	50		

Even if there is heterogeneity of variances within MZ pairs, one can only take practical advantage of genotype–environment interaction or correct the data for its presence if there is a fairly simple relationship between pair sums and differences. Accordingly we obtained the best second-degree polynomial regression curve of differences on sums for each character, although this is not strictly correct since sums as well as differences have variances attached.

The analysis of variance for the second degree polynomial regression for English is shown in Table 4.

Neither the linear nor the quadratic term is significant and this was also true for all the other characters so there is no evidence for these types of interaction in these data.

FITTING MODELS TO THE DATA

We are now in a position to perform the analysis of variance on the data for the different characters and so obtain the between and within pair mean-squares. These are shown for the combined MZ and DZ samples in Table 5, together with the number of pairs on which they are based. Since we cannot pool height data for the two sexes, the meansquares for males and females are shown separately for this character.

We wish to estimate parameters which will adequately account for the observed meansquares by fitting to them models which involve the various parameters we think might be important. To do this we use the method of weighted least squares which has been extensively described in the literature, as have the various parameters and models discussed below (Jinks & Fulker, 1970; Mather & Jinks, 1971; Eaves, 1972; Eaves & Eysenck, 1975*b*). Some of the lower numbers on which meansquares are based in these data are probably pushing the method to the limit of its application.

The full model matrix for the meansquares of twins reared together from a population which is in equilibrium for a given degree of assortative mating is shown in Table 6. It involves within-family (E_1) and between-family (E_2) environmental components, additive (D_R) and dominance (H_R) genetic components and Fisher's assortative mating parameter (A), the correlation between the additive genetical deviations of spouses.

We have shown in the previous paper that only the more intelligent two-thirds of the cohort took the Intermediate examinations so to that extent any estimates of parameters obtained might not be population estimates.

Even if we omit A and assume random mating, the remaining four parameters can still not be estimated because the rows of the model are not independent. However, it is pointless to fit more parameters than are required to account for the observed variation so we shall first fit a single parameter, E_1 model (i.e. setting $D_R = H_R = E_2 = 0$) and then the simple environmental E_1

Table 5. *Number of pairs, between- and within- meansquares for combined MZ and DZ groups for each character*

	MZ			DZ		
	N	Between M.S.	Within M.S.	N	Between M.S.	Within M.S.
English	51	383·74	41·27	79	253·27	116·37
French	13	1054·71	72·92	17	472·37	212·06
History	28	276·29	119·07	36	344·87	187·88
Geography	25	416·63	42·02	35	284·35	150·96
Mathematics 1	31	808·12	63·92	48	379·23	194·92
Mathematics 2	23	726·77	64·74	34	430·88	126·97
Physics	15	674·56	71·23	17	355·85	105·88
Chemistry	15	921·35	46·20	16	707·35	131·09
Science 1	17	652·42	73·94	25	367·30	99·02
Science 2	17	862·12	84·24	25	405·19	160·58
IQ	27	180·27	25·57	19	211·24	90·79
Height (females)	17	143·92	1·59	12	42·98	9·05
Height (males)	10	75·03	2·01	7	83·13	16·09

Table 6. *Full model for meansquares of MZ and DZ twins reared together*

	D_R	H_R	E_1	E_2
MZ between pairs	$1+\frac{A}{1-A}$	$\frac{1}{2}$	1	2
MZ within pairs	0	0	1	0
DZ between pairs	$\frac{3}{4}+\frac{A}{1-A}$	$\frac{5}{16}$	1	2
DZ within pairs	$\frac{1}{4}$	$\frac{3}{16}$	1	0

and E_2 model and simple genetical E_1 and D_R model. Only if both the two parameter models fail shall we be obliged to fit any of the various three parameter models.

For the E_1D_R model we may estimate the heritability as $\frac{1}{2}\hat{D}_R/(\frac{1}{2}\hat{D}_R+\hat{E}_1)$ and calculate its variance as shown by Eaves (1970).

The E_1 model, which tests whether there is variation in addition to sampling error and individual environmental variation, was fitted to all ten statistics (only eight for IQ where no opposite-sex pairs were tested) for the five twin classes combined over years and to the four statistics from the combined MZ and DZ classes. In both cases it failed badly for all 11 characters except History, where it fitted ($P = 0{\cdot}09$) for the ten statistics and only just failed ($P = 0{\cdot}04$) for the four statistics. It appears that most of the variation in performance in the History examination arises from specific environmental influences acting on individuals or from the unreliability inherent in the method of its assessment.

Table 7 shows the simple environmental and genetical models fitted to all ten statistics for English while Table 8 shows the two models fitted to the four statistics from the combined MZ and DZ classes for the same subject.

It can be seen in the case of the ten statistics that even though the deviations are quite large the E_1E_2 model just fits. Because there is little heterogeneity between the classes of different sex, the estimates of the parameters from the four statistics are practically the same but we now find

Table 7. *Simple environmental and simple genetical models fitted to ten statistics for English*

		E_1E_2 model		E_1D_R model	
	Observed M.S.	Weight	Expected M.S.	Weight	Expected M.S.
MZF					
b	369·4323	0·000161	295·1685	0·000122	338·6461
w	44·3103	0·001920	86·9115	0·008540	41·2115
MZM					
b	380·5022	0·000121	295·1685	0·000092	338·6461
w	37·2727	0·001456	86·9115	0·006478	41·2115
DZF					
b	242·4121	0·000132	295·1685	0·000165	264·2875
w	120·5625	0·001589	86·9115	0·000898	115·5702
DZM					
b	195·1786	0·000115	295·1685	0·000143	265·2875
w	109·9762	0·001390	86·9115	0·000786	115·5702
DZOS					
b	275·2233	0·000189	295·1685	0·000236	264·2875
w	117·3676	0·002251	86·9115	0·001273	115·5702

E_1E_2 model: $\chi^2_8 = 15·05$, $p = 0·06$,
$\hat{E} = 86·9115 \pm 10·7800$,
$\hat{E}_2 = 104·1285$, $\pm 19·4307$,

E_1D_R model: $\chi^2_8 = 1·30$, $h^2 = 0·78 \pm 0·05$.
$\hat{E}_1 = 41·2115 \pm 7·8802$,
$\hat{D}_R = 297·4346 \pm 40·6366$.

Table 8. *Simple environmental and simple genetical models fitted to four statistics for English*

		E_1E_2 model		E_1D_R model	
	Observed M.S.	Weight	Estimated M.S.	Weight	Expected M.S.
MZ					
b	383·7416	0·000270	304·2337	0·000212	346·9547
w	41·2745	0·003376	86·9115	0·015107	40·9631
DZ					
b	253·2671	0·000421	304·2337	0·000544	270·4568
w	116·3734	0·005229	86·9115	0·002902	117·4610

E_1E_2 model: $\chi^2_2 = 14·37$ $(P < 0·001)$,
$\hat{E}_1 = 86·9115 \pm 10·7800$,
$\hat{E}_2 = 108·6611 \pm 19·7638$,

E_1D_R model: $\chi^2_2 = 0·45$.
$E_1 = 40·9631 \pm 7·8624$,
$\hat{D}_R = 305·9915 \pm 40·7307$.

that the residual chi square for 2 degrees of freedom clearly discriminates between the E_1E_2 model, which fails badly, and the E_1D_R model, which gives an excellent fit.

The final chi squares for the two models fitted to both ten and four statistics are shown for all 11 characters in Table 9. In each case these stable values were obtained after about four iterations. The heritability and its standard error estimated from the E_1D_R fit to the four statistics is also given for each character.

The poor resolution of model fitting to ten statistics is confirmed, both models fitting at the 5 % significance level in most cases. The heterogeneity of meansquares from the different sexes can be gauged by comparing values of χ^2_8 with their corresponding χ^2_2 values obtained by fitting the same model to the four statistics from the combined data. There is particularly marked heterogeneity for Science 2, where both χ^2_8 values are significant and it is also evident for History, Chemistry and Science 1. If sample sizes were larger it would be interesting to fit different parameters for the two sexes in these subjects.

Table 9. *Final chi squares and heritabilities for simple environmental and genetical models fitted to ten*[a] *and four statistics*

	χ_8^{2a} values		χ_2^2 values		
	E_1E_2	E_1D_R	E_1E_2	E_1D_R	Heritability
English	15·05†	1·30	14·37***	0·45	0·79 ± 0·05
French	8·40	3·43	5·33†	1·36	0·83 ± 0·07
History	9·79	11·51	1·86	1·34	0·47 ± 0·13
Geography	14·61†	4·47	8·84*	0·47	0·81 ± 0·06
Mathematics 1	15·31†	4·12	13·49**	2·53	0·81 ± 0·05
Mathematics 2	7·71	6·40	4·48	2·82	0·81 ± 0·06
Physics	3·34	3·80	2·08	2·41	0·77 ± 0·09
Chemistry	9·30	7·86	3·70	2·00	0·89 ± 0·05
Science 1	7·89	9·22	2·09	3·54	0·76 ± 0·09
Science 2	19·52*	16·99*	4·65	2·50	0·77 ± 0·08
IQ	18·44**[a]	8·62[a]	8·74*	1·20	0·79 ± 0·06

[a] No opposite sex pairs for IQ so only eight statistics and χ_6^2 values.
† $0·05 < P < 0·10$. * $0·01 < P < 0·05$. ** $0·001 < P < 0·01$. *** $P < 0·001$.

When we turn to the models fitted to four statistics we find that the simple environmental model fails badly for English, Mathematics 1, Geography and IQ and almost fails ($P = 0·07$) for French. On the other hand, the simple genetical model fits well for all characters.

Thus, for IQ and four of the ten examination subjects, a simple environmental model is insufficient to explain the observed variation while a simple genetical model will. In the other six subjects the numbers are not large enough to exclude either hypothesis. The heritabilities for all 11 characters except History, which we have discussed, lie between 0·76 and 0·89 with fairly small standard errors, although these, of course, only refer to repeat studies of the same tests in the same population.

In the case of height, because the means for males and females are different we must fit models separately to data for the two sexes. However, we observe that the total variance for MZ females is significantly greater ($P < 0·001$) than that for DZ females and this will cause failure of any simple model, which assumes variances to be equal. On the other hand, the variances for males alone are based on such small numbers that the model fitting procedure becomes invalid. We must conclude that sampling with respect to height is inadequate to allow estimation of population parameters. However, we can observe that the DZ within-pair meansquare is much greater than the corresponding MZ meansquare for both males and females and this is evidence for a strong genetic component in phenotypic variation for height. The analysis illustrates the importance of working with meansquares where intraclass correlations would obscure such inadequacies of sampling.

HOW MANY DIFFERENT GENETIC COMPONENTS?

Are the significant heritable components for the different subjects all measures of the same well-documented general intelligence component, or does performance in the different subjects involve specific genetic components in addition to general ability?

Eaves & Gale (1974) review earlier attempts to answer this question and show how the biometrical genetical approach can be extended to detect and elucidate the types of gene action involved in genetical covariation. However, all previous methods rely on the fact that all twins

Table 10. *Analyses of variance for MZ and DZ twins of traits (English and Mathematics 1) cross classified with pairs*

	MZ		DZ	
	D.F.	M.S.	D.F.	M.S.
Between pairs	30	3·2612	47	1·7028
Within pairs	31	0·2595	48	0·9857
Traits	1	16·3054	1	23·9974
Between-pairs × traits	30	1·2945	47	0·6633
Within-pairs × traits	31	0·1215	48	0·3136

took all subtests. For this twin sample, the numbers of individuals who take even three subjects in common is not large so that the possibilities for multi-variate analysis are fairly limited. As a first attempt though, we can try to detect the effects of specific genes on specific traits by seeking evidence of heterogeneity of gene action on the different traits using a method that has been discussed by Wilson (1968). This entails detecting a genetical component in the interaction meansquares obtained by cross-classifying twin pairs with traits in an analysis of variance design.

Any heterogeneity so detected does not necessarily imply the existence of two sets of loci acting independently. Eaves & Gale (1974) have pointed out that it could also represent one set of genes affecting both traits in the same direction and a second bipolar set in which increasing alleles for one trait are decreasing for the other. Furthermore, lack of heterogeneity does not necessarily indicate pleiotropic effects of a single set of loci but could represent more than one set in marked linkage disequilibrium.

The search for heterogeneity was restricted to pairwise comparisons among the six characters which had shown definite evidence of genetic components. Only those comparisons involving at least ten pairs of each twin type were made. Since we are only interested in deciding whether genetical covariance between characters is not complete, the sampling difficulties associated with height do not prevent us from making comparisons with this character.

It would be possible to detect the interactions we seek merely as an artefact of the different total variances of the two characters being compared, particularly in comparison involving IQ and Height which we know to have different variances. Accordingly, we converted raw scores for each character to standard scores by dividing by the appropriate standard deviation calculated from the total twin sample. We then checked by Bartlett's test that the four total variances of the standardized scores (two characters for MZ and DZ) were homogeneous.

Table 10 shows the analysis of variance of standardised scores for MZ and DZ twins with individuals crossclassified with traits and grouped hierarchically into pairs in this case for the two subjects English and Mathematics 1.

We can now use the between-pairs × traits and within-pairs × traits meansquares for MZ and DZ twins as our four meansquares for model fitting as before. The E_1 model failed for all ten comparisons and the final chi squares after fitting the simple environmental and genetical models are shown in Table 11.

The E_1E_2 model fails at the 5 % level in the English–Mathematics 1, English–Height and IQ–Height comparisons while the E_1D_R model fits all comparisons. Because the comparisons are not independent of one another, the overall significance of the results might be somewhat overestimated. However, in this case different twin pairs are involved in each comparison, only a

Table 11. *Simple environmental and genetical models fitted to interaction meansquares for standardized scores of different pairs of characters*

(Numbers of MZ and DZ pairs in each comparison are also shown.)

			χ^2_2 values	
	N_{MZ}	N_{DZ}	E_1E_2	E_1D_R
English–French	13	17	1·50	4·02
English–Geography	25	34	5·82†	2·31
English-Mathematics 1	31	48	10·54**	1·98
English–IQ	27	19	4·55	0·09
English–Height	27	19	8·08*	3·01
French–Mathematics 1	12	15	3·47	3·88
Geography–Mathematics 1	13	19	0·03	1·48
Mathematics 1–IQ	15	12	1·85	2·34
Mathematics 1–Height	15	12	1·31	5·73†
IQ–height	27	19	9·48**	1·81

Symbols as for previous tables.

small number being common to all of them and this tends to increase the independence of the tests.

We may tentatively conclude that there are differences in the action of genes within the pairs of traits mentioned. Taking account of the numbers in each comparison it seems very reasonable that in two of the three comparisons with Height and in the English–Mathematics comparison where one would most expect to find it, genetical heterogeneity is found. But in the interactions of examination subjects with IQ and of English and French where one would least expect to find it with such low powered tests as these, no such heterogeneity is detected.

DISCUSSION

It has been suggested by Jensen (1969) and repeated by others that the heritability of educational achievement is considerably lower than that for IQ. Educational achievement tests by their nature are much more heterogeneous than IQ tests and it is not surprising to find some low heritabilities reported. Nevertheless, some of the Stanford Achievement tests of Newman, Freeman & Holzinger (1937) and others of Husén (1960) show quite large genetic components of variation while Jinks & Fulker (1970) have found serious inadequacies in the oft-quoted low heritabilities from Burt's (1966) educational achievement data. More recently Nichols (1965) has found most heritabilities for the different subtests of the National Merit Scholarship Qualifying Test in the 0·6–0·7 range as have Trimble & Mi (1973) for SCAT and STEP achievement tests in Hawaii. The results of this study support the view that educational achievements, at least at the level of large scale impersonal testing programs, are not much less heritable than IQ; indeed in this study they are in the same range.

One factor which may considerably lower heritabilities of educational achievements is the low reliability of the test scores. Eaves & Eysenck (1975*a*) have recently suggested that test–retest reliability may itself have a genetical component, indicating genetic variability in unexplored areas of human personality. However, in subjects such as English where subjective judgement is so important in the assessment it is quite likely that simple measurement error comprises a considerable proportion of the non-heritable variation.

The fact that a simple genetical model fits most of the characters in these data does not exclude the existence of effects due to environmental differences between families, to dominance or assortative mating, but shows that if they do exist they are contributing insufficiently to the observed variation to be detected. In fact Eaves (1973, 1975), has found evidence that all these factors contribute to variation in IQ but he has also shown (Eaves, 1972) that large sample sizes of more efficient data sets than are available here are required to reliably detect even large amounts of dominance.

The controversy over the number and nature of different factors influencing human abilities has been a long one in psychology and has been discussed by Cattell (1971). Nichols found evidence for heritable specific cognitive abilities from his large study of twins who took the National Merit Scholarships Qualifying Test and other evidence has been reviewed by Vandenberg (1968) and Mittler (1971). Bock & Kolakowski (1973) have recently presented evidence for a sex-linked gene affecting 'spatial visualisation'.

From this study we can only say that we have evidence for some genes affecting English differently from Mathematics and Height, and IQ differently from Height. We should like to be able to state how much of the variance in, say, English is due to genes commonly affecting performance in all subjects and how much is due to genes affecting English only. Eaves & Gale (1974) have made a start on this problem and it is hoped soon to publish improvements to their methods.

SUMMARY

1. Examination results for 98 pairs of same-sex and 39 pairs of opposite-sex twins were available for analysis as well as measurements of IQ and height for 47 of the same-sex pairs.

2. The choice of subjects taken for examination is not influenced by genetic differences within DZ pairs but seems to be mainly determined by between family cultural differences.

3. Neither the sex difference in opposite-sex pairs nor the difference between first and second born twins in same-sex pairs contributes appreciably to within pair environmental variance. However, same-sex DZ pairs are more often found in different school classes than MZ pairs.

4. No evidence for interaction between genotypes and within-family environments was found for any character.

5. The simple environmental model fails to fit the data for English, French, Geography, Mathematics and IQ whereas the simple genetical model fits. For the remaining subjects it is not possible to exclude either hypothesis. The heritabilities of the examination performances are of the same magnitude as that for IQ.

6. Sampling inadequacies are revealed in the data for height but there is evidence for a substantial genetic component of phenotypic variation.

7. There is evidence for heterogeneity of gene action in the subjects English and Mathematics supporting the view that there are genetically determined specific abilities acting independently of general intelligence.

I should like to thank Dr L. J. Eaves for his helpful criticism and suggestions during the analysis of these data and Professor J. L. Jinks for his critical reading of the manuscript.

REFERENCES

BOCK, R. D. & KOLAKOWSKI, D. (1973). Further evidence of sex-linked major gene influence on human spatial visualizing ability. *American Journal of Human Genetics* **25**, 1–14.

BURT, C. (1966). The genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart. *British Journal of Psychology* **57**, 137–53.

CATTELL, R. B. (1971). *Abilities: Their Structure Growth and Action.* Boston: Houghton Mifflin.

EAVES, L. J. (1970). The genetic analysis of continuous variation: a comparison of designs applicable to human data. II. Estimation of heritability and comparison of environmental components. *British Journal of Mathematical and Statistical Psychology* **23**, 189–98.

EAVES, L. J. (1972). Computer simulation of sample size and experimental design in human psychogenetics. *Psychological Bulletin* **77**, 144–52.

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

EAVES, L. J. (1975). Testing models for variation in intelligence. *Heredity* **34**, 132–6.

EAVES, L. J. & EYSENCK, H. J. (1975*a*). Genetical and environmental components of inconsistency and unrepeatability in twins' responses to a neuroticism questionnaire. *Behaviour Genetics* (in the Press).

EAVES, L. J. & EYSENCK, H. J. (1975*b*). The nature of extraversion: a genetical analysis. *Journal of Personality and Social Psychology* (in the Press).

EAVES, L. J. & GALE, J. S. (1974). A method for analysing the genetical basis of covariation. *Behaviour Genetics* **4**, 253–67.

HUSÉN, T. (1960). Abilities of twins. *Scandinavian Journal of Psychology* **1**, 125–135.

JENSEN, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review.* **39**, 1–123.

JINKS, J. L. & FULKER, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behaviour. *Psychological Bulletin* **73**, 311–49.

MARTIN, N. G. & MARTIN, P. G. (1975). The inheritance of scholastic abilities in a sample of twins. *Annals of Human Genetics* **39**, 213–8.

MATHER, K. & JINKS, J. L. (1971). *Biometrical Genetics*, 2nd ed. London: Chapman and Hall.

MITTLER, P. (1971). *The Study of Twins.* London: Penguin.

NEWMAN, H. H., FREEMAN, F. N. & HOLZINGER, K. J. (1937). *Twins: a Study of Heredity and Environment.* Chicago: University of Chicago Press.

NICHOLS, R. C. (1965). The national Merit Twin Study. In *Methods and Goals in Human Behaviour Genetics.* (ed. S. G. Vandenberg). New York: Academic Press.

TRIMBLE, B. K. & MI, M. P. (1973). Estimates of components of variation in academic aptitude and achievement scores of students in Hawaii. *Human Heredity* **23**, 501–10.

VANDENBERG, S. G. (1968). The nature and nurture of intelligence. In *Genetics.* (ed. D. C. Glass). New York: Rockefeller University Press.

WILSON, R. S. (1968). Autonomic research with twins: methods of analysis. In *Progress in Human Behaviour Genetics* (ed. S. G. Vandenberg). Baltimore: Johns Hopkins Press.