THE GENETIC STRUCTURE OF PERSONALITY—II.
GENETIC ITEM ANALYSIS OF THE EPQ

A. C. HEATH, R. JARDINE,*† L. J. EAVES and N. G. MARTIN‡
Department of Human Genetics, Box 33, Medical College of Virginia,
Virginia Commonwealth University, Richmond, VA 23298, U.S.A.

(Received 13 June 1988)

Summary—Most studies of the inheritance of personality have performed genetic analysis of scale scores, which will sometimes give a misleading impression of what is inherited. Item genetic analysis, where this is focused on differences in genetic architecture rather than merely differences in item heritability, should guide the development of genetically homogeneous scales. Item genetic analyses of the EPQ were performed, using data from 3810 pairs from the Australian Twin Register. All but three items exhibited significant genetic variance. Genetic dominance was found for a majority of items from the Extraversion and Neuroticism scales, although the dominance effects were weak for N items, particularly in females. Shared environmental effects were found for items from the Lie and Psychoticism scales; for P items they were strong in males but weak in females. These differences in genetic architecture and genotype × sex interaction should permit use of multivariate item genetic analysis to identify the underlying genetic structure of the personality dimensions assessed by the EPQ.

INTRODUCTION

Studies of twins reared together, separated twin studies, and adoption studies, give consistent evidence for a substantial influence of heredity on major dimensions of personality (see reviews by Eaves and Young, 1981; Fulker, 1981; Eaves, Eysenck and Martin, 1989). Most published genetic research, however, has focused upon the analysis of scale scores, which have been derived from factor analyses of personality data from samples of unrelated individuals. Recent multivariate genetic analyses of twin personality and symptom data (Kendler, Heath, Martin and Eaves, 1987; Silberg, Martin and Heath, 1987; Martin, Jardine, Andrews and Heath, 1988; Heath and Martin, unpublished) have provided examples where the observed phenotypic factor structure is determined by quite different underlying genetic and environmental factor structures. Under these conditions a quite misleading notion of what it is that is inherited can arise from conventional genetic analyses. Kendler et al. (1987), for example, were able to confirm by conventional factor analysis of self-report symptom data the existence of two separate anxiety and depression phenotypic factors; but used multivariate genetic analysis to show that a single genetic factor explained most of the common genetic variance in the symptoms. In the first paper in this series (Heath, Jardine, Eaves and Martin, 1988), we were able to replicate the phenotypic factor structure predicted for the EPQ (Eysenck and Eysenck, 1975, 1976) in a four-factor solution; but multivariate genetic analysis of the P-scale items (Heath and Martin, unpublished) revealed a dimension of heritable variation quite different from that predicted.

Genetic item analysis has been advocated as one means of using genetic information to guide the construction of scales. Most published analyses, however, have focused on differences in item heritability (Horn, Plomin and Rosenman, 1976; Loehlin, 1965, 1986; Loehlin and Nichols, 1976; Neale, Rushton and Fulker, 1986). These can arise through differences in item reliability, differences in item loadings on an underlying genetic common factor, or differences in item-specific genetic influences. In the previous paper in this series (Heath et al., 1988), we have suggested that item differences in ‘genetic architecture’ (Jinks and Fulker, 1970; Eaves and Young, 1981; Fulker, 1981) may offer a surer foundation for uncovering the genetic structure of personality. By exploiting the large sample sizes now being used in twin research (e.g. Martin and Jardine, 1986;
Floderus-Myrhed, Pedersen and Rasmuson, 1980; Rose, Koskenvuo, Kaprio, Sarna and Langinvainio, 1988), it might be possible to uncover differences in mode of inheritance between different personality dimensions. One dimension may exhibit genetic dominance, a second genotype $\times$ sex interaction, a third purely additive gene action with no sex interaction, and so forth. Such differences in genetic architecture would allow us to address issues of factor rotation and the choice between primary vs higher-order factor solutions (e.g. Eysenck, 1981; Gray, 1981) which remain intractable in conventional factor analytic studies. In this second paper, we explore the genetic architecture of the EPQ using univariate item genetic analyses of Australian twin data (Martin and Jardine, 1986).

METHOD

Sample

A health questionnaire, incorporating the full Eysenck Personality Questionnaire (Eysenck and Eysenck, 1975), was mailed to adult twin pairs aged 18 years and older who had registered with the Australian NH&MRC volunteer Twin Register. Completed questionnaires were received from both members of 3810 pairs, 64% of the 5967 pairs to whom questionnaires were mailed. These included 1233 monozygotic female, 567 monozygotic male, 751 dizygotic female, 352 dizygotic male and 907 unlike-sex pairs. Zygosity was determined on the basis of responses to a brief two-item questionnaire, a method of assessment which has been found to give a misclassification rate of less than 5% when compared to the results of blood-typing (e.g. Nichols and Bilbro, 1966; Martin and Martin, 1975; Kasriel and Eaves, 1976). A subsample of 96 of the twins had previously completed and returned a pilot version of the same questionnaire, on average three months beforehand. For these individuals, test-retest data was available for the EPQ items.

Data-summary

For each EPQ item, two-way contingency tables were computed for each of the five twin zygosity groups, cross-classifying response of first twin by response of second twin. Twins from like-sex pairs were designated as first or second twins on the basis of birth-order, where such information was available, but otherwise at random. For twins from unlike-sex pairs, responses of the male twin were cross-classified by responses of the female twin from each pair. For the reliability subsample, two-way tables were computed, cross-classifying response on the first occasion by response on the second occasion. For this subsample, the number of observations was too small to permit estimation of separate reliability coefficients for the two sexes.

Tetrachoric correlations were estimated for each contingency table by the method of maximum-likelihood (Olsson, 1979; Joreskorg and Sorbom, 1985). Estimation of the tetrachoric correlation implies the assumption that the discontinuous response distribution observed for a given item is determined by an underlying latent personality variable whose distribution is continuous and normal. A threshold divides this continuous latent distribution into the observed dichotomous distribution. It is further assumed that for pairs of responses to an item (by two members of a twin pair, or by the same individual from the reliability sample on two occasions), the joint distribution of the underlying latent personality variables is bivariate normal. With items which are dichotomous, we have no statistical test of these assumptions, but for personality items they appear appropriate. The log-likelihood of a contingency table, with observed frequency $f_{ij}$ for the $i, j$th cell, and corresponding expected frequency $p_{ij}$ under a given model, will be

$$L = \ln(c) + \sum f_{ij} \ln(p_{ij})$$

(1)

where $c$ is a constant. Under the threshold model, with threshold values $t_i$ and $t_j$ for the first and second twins (or occasions), and correlation between latent personality variables in the first and second twins $r$, the expected cell frequencies will be given by

$$p_{ij} = \Phi(t_i, t_j) - \Phi(t_{i-1}, t_j) - \Phi(t_i, t_{j-1}) + \Phi(t_{i-1}, t_{j-1}),$$

(2)

where $\Phi$ is the bivariate normal distribution function with correlation $r$, $t_0 = t'_0 = -\infty$, and $t_2 = t'_2 = +\infty$. Thus maximum-likelihood estimates of the correlation $r$, and threshold values $t_i$ and
The genetic structure of personality—II 617
ti, are obtained by maximizing the log-likelihood in (1) with respect to these parameters. Except in the case of unlike-sex pairs, where allowance must be made for sex-differences in item endorsement frequency, we would expect to find that ti = ti, i.e. that threshold values are the same for first and second twins, for any given item. Tetrachoric correlations were estimated using a user-written program, since commercially available software (e.g. Joreskog and Sorbom, 1985) does not permit threshold values to be constrained to be equal for first and second twins from like-sex pairs, nor for different zygosity groups (see below).

**Estimation of variance components**

By a simple extension of these methods, maximum-likelihood estimates of genetic and environmental components of variance can be obtained by fitting models to contingency tables for monozygotic and dizygotic twin pairs (Eaves, Last, Young and Martin, 1978). Ignoring, for simplicity, genotype × sex interaction, the expected correlation between monozygotic twin pairs will be VA + VD + EC, and that between dizygotic twin pairs will be 1/2 VA + 1/4 VD + EC, where VA is the additive genetic component of variance, VD is the dominance genetic component of variance, and EC is the shared environmental component of variance. The component of variance due to individual-specific environmental effects (i.e. those environmental effects not shared by members of a twin pair) is obtained by subtraction as ES = 1 - VA - VD - EC. For any given values of our genetic and environmental parameters, and thresholds (which will be allowed to differ as a function of sex but not zygosity), we can compute the log-likelihood of an observed contingency table by substituting in (2) and (1) above. The overall log-likelihood of a set of observed tables is obtained by summing across tables. Thus maximum-likelihood estimates of variance components and threshold values were obtained by maximizing the overall log-likelihood with respect to the genetic and environmental parameters and threshold values.

Without separated twin data, the effects of dominance and shared environment will be confounded, strong effects of genetic dominance leading to negative estimates of the shared environmental component of variance, and vice versa (e.g. Martin, Eaves, Kearsley and Davies, 1978). Thus if we estimate parameters VA and EC and obtain a negative estimate of EC, this indicates that we should be estimating VA and VD, and vice versa. Strong effects of genetic dominance will be needed, however, to mask shared environmental influences in twin data (Martin et al., 1978). We therefore began by estimating VA and EC rather than VA and VD.

In most genetic applications of model-fitting, we are seeking the simplest model which is consistent with the observed data, i.e. gives an adequate fit to the data by chi-square test of goodness-of-fit, and does not give a significantly worse fit than more elaborate models by likelihood-ratio χ² test (Eaves et al., 1978). For the genetic item analysis of a 90-item questionnaire, however, this approach is impractical and, arguably, undesirable. In the present case we wished to compare the genetic architectures of items from the E, N, P and L scales of the EPQ to determine if consistent differences between scales could be detected. From the work of Martin et al. (1978) it is known that the power of the classical twin study to resolve additive genetic and shared environmental effects increases with decreasing specific environmental variance (which will include variance due to measurement error). This will also be the case for the resolution of additive genetic and dominance effects, and for the detection of sex-differences in genetic and shared environmental variance components. In conventional model-fitting, estimates of genetic dominance would be reported only if the VA VD model gave a significantly better fit than the VA model alone. In our application, however, this decision rule could give rise to apparent differences in genetic architecture which are solely a function of differences in item reliability between scales.

We therefore adopted a different approach. Initially we estimated variance components VA and EC from the set of 4 like-sex contingency tables for each item, ignoring any sex differences in these parameters. Unlike-sex twin pairs were excluded from the analysis because genotype × sex interaction, such that the correlation between gene effects or between shared environmental effects in the two sexes is less than unity, can give rise to unlike-sex dizygotic correlations which are lower than the corresponding same-sex dizygotic correlations. If such sex interactions are ignored, i.e. dizygotic twin correlations are in effect pooled, this will tend to lower the dizygotic correlation and hence mimic the effects of genetic dominance. Estimates of EC and (VA + EC) were compared to provide an index, which could be compared across items, of what proportion of the familial
variation—i.e. ignoring measurement error plus specific environmental effects—was attributable to shared environmental effects (positive values of the ratio) or genetic dominance (negative values). In further analyses, again ignoring the unlike-sex pairs, we obtained estimates of this ratio separately for male and for female like-sex pairs.

RESULTS

For each item, Table 1 gives estimates of polychoric correlations for each twin group, and for the reliability sample. For a few items from the P-scale, the item endorsement frequency was too low to permit estimation of the test–retest correlation. Reliability coefficients were often substantially higher than the corresponding monozygotic twin correlations. This indicates that some of the individual-specific environmental variance reflects factors which are relatively stable over time (at least over a three month test–retest interval). Sample sizes for the pilot sample were small, however, and the correlations given are inflated by the pooling of male and female respondents. Monozygotic twin correlations for many items were strikingly high, indicating the importance of familial influences even at the item level (cf. Martin, Eaves, Heath, Jardine, Feingold and Eysenck, 1986).

Figures 1–4 plot estimates of EC against (VA + EC) from the analyses ignoring sex differences, separately for items from the Extraversion, Neuroticism, Lie and Psychoticism scales. In all but 3 items (P2, P9, L7) positive estimates of the additive genetic component of variance of 0.1 or greater were obtained, so we focus our attention on differences in genetic architecture. For Extraversion, only 4/21 items gave positive estimates of EC, and 14/21 gave negative estimates, consistent with findings of significant genetic dominance for the scale scores for this variable (Martin and Jardine, 1986). For Neuroticism, 9/23 items gave positive estimates of EC, and 12/23 items gave negative estimates of EC. For this trait, no significant evidence for either genetic dominance or shared environmental effects was found in the analysis of scale scores (Martin and Jardine, 1986), although in subsequent multivariate analysis Martin et al. (1988) have found evidence for dominance effects on male N scores. For the Lie scale (18/21 items with positive estimates of EC), consistent evidence for shared environmental effects was obtained. For Psychoticism, 16/25 items gave positive estimates of EC, but some items gave strongly negative estimates implying genetic dominance. Means of the estimated ratios EC/(VA + EC) were −0.20 for Extraversion, −0.05 for Neuroticism, 0.23 for Psychoticism and 0.25 for Lie items. The presence or absence of genetic dominance appears to discriminate Extraversion from Neuroticism items, but shared environmental effects are important for both Psychoticism and Lie items.

Estimates of genetic and shared environmental variance components for males and females, from the analyses which allowed for sex-dependent parameters, are given in Table 1. Where negative estimates of EC were obtained, the VA VD model was also fitted, and estimates of these components have been tabulated instead. Overall item broad heritabilities, for comparison with

Table 1. Estimates (×100) of twin tetrachoric correlations and additive genetic (VA), dominance genetic (VD) and shared environmental (EC) components of variance, for EPQ items grouped by scale. († indicates correlation could not be reliably estimated because of low endorsement frequencies)

<table>
<thead>
<tr>
<th>EPQ items</th>
<th>Twin polychoric correlations</th>
<th>Estimates of variance components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MZF</td>
<td>MZM</td>
</tr>
<tr>
<td>1 E1</td>
<td>51</td>
<td>40</td>
</tr>
<tr>
<td>5 E2</td>
<td>57</td>
<td>49</td>
</tr>
<tr>
<td>10 E3</td>
<td>59</td>
<td>54</td>
</tr>
<tr>
<td>14 E4</td>
<td>55</td>
<td>63</td>
</tr>
<tr>
<td>17 E5</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>21 E6</td>
<td>52</td>
<td>45</td>
</tr>
<tr>
<td>25 E7</td>
<td>47</td>
<td>31</td>
</tr>
<tr>
<td>29 E8</td>
<td>48</td>
<td>42</td>
</tr>
<tr>
<td>32 E9</td>
<td>57</td>
<td>41</td>
</tr>
<tr>
<td>36 E10</td>
<td>51</td>
<td>37</td>
</tr>
<tr>
<td>40 E11</td>
<td>49</td>
<td>34</td>
</tr>
<tr>
<td>42 E12</td>
<td>57</td>
<td>49</td>
</tr>
<tr>
<td>45 E13</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>49 E14</td>
<td>59</td>
<td>52</td>
</tr>
<tr>
<td>52 E15</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>56 E16</td>
<td>36</td>
<td>28</td>
</tr>
</tbody>
</table>
results reported by previous studies, are given by VA (where positive or zero estimates of EC were obtained), or VA + VD if dominance was found. For some items negative estimates of the additive genetic component have been obtained under the VA VD model, reflecting the strong negative correlation between estimates of VA and VD which is found in analyses of twin data. Estimates of the broad heritability will still be unbiased in these cases, but we cannot apportion the contribution of additive and dominant gene action with any confidence. In a few cases, the estimate of a parameter has gone to the upper bound of unity. In such cases the broad heritability estimate will not be unbiased.

As in the analyses ignoring sex-differences, positive estimates of the shared environmental component were obtained for the majority of items from the Lie scale (15/21 for males, 18/21 for females, including respectively 10, 12 items with EC > 0.10) and from the Psychoticism scale (17/25 for males, 16/25 for females with 16, 15 items with EC > 0.10). For Neuroticism, 13/23 items in males and 11/23 items in females gave positive estimates of genetic dominance (including 10 items in each sex with VD > 0.10), compared with 12/21 items (males) and 14/21 items (females) in the case of the Extraversion items (with 8, 13 items with VD > 0.10). This comparison does not suggest a clear discrimination between Neuroticism and Extraversion items on the basis of genetic architecture. However, Table 2 gives mean values of the ratio EC/(VA + EC) for each scale, separately for males and females. Consistent evidence for dominance for Extraversion was found in both sexes. For Neuroticism, however, evidence for dominance is weak in both sexes. Evidence
The genetic structure of personality—II

for shared environmental effects for Lie scale scores is consistent in both sexes; but for Psychoticism items shared environmental effects are weak in females.

DISCUSSION

We have suggested (Heath et al., 1988) that it might be possible to discriminate between the personality dimensions of the EPQ on the basis of genetic architecture. The results of the item genetic analyses reported here appear to confirm that contention. For Extraversion items, we found both additive gene action and dominance. Neuroticism items appeared to show purely additive genetic inheritance. Items from both the Lie scale and the Psychoticism scale were influenced by shared environmental effects, but for the Psychoticism items these shared environmental effects appeared to be largely restricted to males. Inspection of the estimates of genetic and environmental components for individual items (Table I) does not reveal consistent differences in genetic architecture between potential subscales which might be derived from the E, N, L and P scales. Although we found that items from the P scale are rather heterogeneous in their mode of inheritance, with some items showing strong genetic dominance and others shared environmental effects, we were unable to identify differences in architecture between the sub-factors of the P scale reported by Heath et al. (1988).
Fig. 3. Genetic architecture of Psychoticism items (ignoring sex differences). For each item from the Psychoticism scale, the estimate of EC is plotted against (VA + EC). Numbers identify items within the Psychoticism scale. See Table I for correspondence with EPQ item numbers.

The analyses which we have presented in this paper are univariate genetic analyses. They do not use information about the genetic and environmental sources of covariation between different items within a scale. In discussion of the limitations of analyses of item heritabilities, we noted that differences in overall heritability might reflect differences in error variance, differences in loadings on a common genetic factor, or differences in item-specific genetic influences. To avoid confounding effects of genetic architecture with differences in item reliability, we compared the results of fitting the same model to every item. If we are attempting a post-hoc classification of items on the basis of genetic architecture, there would still be the danger that observed differences in genetic architecture reflected item-specific rather than common factor influences. It seems unlikely that we would observe differences in genetic architecture for items from the a priori E, N, L and P scales.

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>-0.20</td>
<td>-0.16</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>-0.03</td>
<td>-0.07</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>0.03</td>
<td>0.60</td>
</tr>
<tr>
<td>Lie</td>
<td>0.30</td>
<td>0.25</td>
</tr>
</tbody>
</table>
if these were due to item-specific effects only. To be confident about this conclusion, however, we need to conduct a multivariate item genetic analysis. This we will address in a subsequent paper (Heath, Eaves and Martin, in press).

Acknowledgements—Data-analysis for this paper was supported by NIH grants GM30250 and AG04934, and by ADAMHA grants AA06781 and MH40828. Many of the analyses were performed by RJ, and presented initially in her PhD thesis. Data collection was funded by grants to NGM, Dr J. D. Mathews and Dr J. B. Gibson from the Australian National Health and Medical Research Council. We acknowledge the roles of Dr Mathews in the establishment of the Australian Twin Register, and of Marilyn Olsen in preparation of the data.

REFERENCES

A. C. Heath et al.


