THE EFFECT OF BIAS ON THE ESTIMATES OF THE GENETIC AND ENVIRONMENTAL INFLUENCES ON RATINGS OF DEPRESSIVE SYMPTOMS

J. L. SILBERG,* J. M. MEYER,† L. J. EAVES,* M. C. NEALE,‡ J. K. HEWITT,* A. C. HEATH,§ AND K. S. KENDLER*†

*Department of Human Genetics, Medical College of Virginia, Box 3, Richmond, Virginia, USA
†Department of Psychiatry, Medical College of Virginia, Box 710, Richmond, Virginia 23298, USA.
‡Department of Psychology, Virginia Commonwealth University, Richmond, Virginia 23284, USA.
§Department of Psychiatry, Washington University Medical School, 4940 Audubon Avenue, St. Louis, Missouri 63110, USA.

SUMMARY
Family history studies in psychiatric genetics often make use of ratings from relatives for diagnosing individuals. To study the extent to which the validity of psychiatric ratings may be systematically affected by characteristics of the informant, as well as the context in which informants are asked to rate their subjects, a structural equation model was fit to the self-ratings of lifetime prevalence of depressive symptoms of 463 monozygotic and 308 dizygotic female twins and the ratings of them by their cotwins. The model estimates genetic, shared environmental, and non-shared environmental components of variance for ratings of depressive symptoms, and corrects for two sources of bias: (1) rater bias, or the extent to which characteristics of the informant systematically affects the rating of similar characteristics in the subject they are rating, and (2) correlated error of measurement, a type of bias that can result from a style of questioning that exaggerates an informant's misperception of similarity between herself and her subject. The fit of the model was excellent ($\chi^2 = 58.96, p = 0.62$), and suggests that the validity of psychiatric information obtained from family members may be significantly affected by questionnaire format as well as twins' tendency to bias or 'project' their own experience of depressive symptoms when rating the cotwins' history of symptoms. Consequently, we have shown that using an informant's rating can lead to twin correlations nearly two times higher than those derived using the twins' self-ratings only, thereby affecting the estimates of genetic and environmental influences on depressive symptoms. The implications of our model for questionnaire design, data collection and analyses that depend upon ratings from relatives for obtaining diagnostic information are discussed.

KEY WORDS— Rater bias, family history method, family study method, twins, depression.

INTRODUCTION
The estimation of genetic and environmental effects on psychiatric illness is founded upon obtaining both accurate and reliable diagnostic information on ill individuals and their families. To this end, two approaches to data collection have traditionally been used. In the family history method (Andreasen et al., 1977), diagnostic information about an individual is obtained from interviews with one or more of the individual's relatives. In contrast, in the family study method, all family members are interviewed directly.

Requests for reprints should be addressed to J. L. Silberg

1049-8931/91/020059-09$05.00
© 1991 by John Wiley & Sons, Ltd.

Received 3 April 1991
Accepted 3 April 1991

Because of the time and expense involved in interviewing all relatives and the unavailability of certain family members, the family history method has appealed to many investigators. However, relying upon indirect diagnostic information from relatives has been shown to lead to serious under-reporting of psychiatric illness (Andreasen et al., 1977; Rimmer and Chambers, 1969; Winokur et al., 1969; Mendlewicz et al., 1975).

When assessing the reliability of the family history method, investigators have typically relied on information derived from direct interview as the criterion against which an informant's rating is evaluated. Efforts at improving the reliability of the family history method have focused primarily on increasing its 'sensitivity', i.e., the ratio of the
number of individuals diagnosed as ill based upon informant’s report, to those similarly diagnosed by direct interview, and to a lesser extent, its ‘specificity’, i.e., the ratio of the number of individuals who are reported as well by others to those who are diagnosed as well through interview. It has been shown that the sensitivity of the family history method can vary depending upon the diagnosis of the subject (Thompson et al., 1982), the number of informants used (Andreasen et al., 1977; Thompson et al., 1982), the relationship between the subject and the informant (Mendlewicz et al., 1975; Thompson et al., 1982), and the stringency of criteria used for diagnosing illness (Andreasen et al., 1977). Despite attempts at improving the reliability of the family history method, sensitivity has generally been low, ranging from only 30% to 60%.

Although investigators have studied a number of factors that can affect the reliability of diagnostic information obtained using the family history method, little attention has been paid to the validity of this information, that is, the degree to which psychiatric ratings of the subject may be systematically affected by the phenotype or characteristics of the person doing the rating. To examine this, we introduce the concept of ‘rater bias’ which we define as the extent to which particular features of the informant influence the ratings of similar characteristics of the subject. We also consider whether the validity of psychiatric ratings may be affected by a general misperception of similarity among relatives and a questionnaire or interview format that contributes to this misperception. This might occur if an informant perceives herself to be similar to her subject in a variety of ways (not necessarily related to the specific behaviours that she is asked to rate), and simply ascribes the same pattern of responses to the subject as she has given to herself. Examples of the kind of questionnaire format that could potentially exaggerate this perception of similarity is one that asks the subject to rate herself and her subject on the same occasion and/or has both sets of responses readily available for comparison.

To explore rater bias and the effect of questionnaire format on the validity of psychiatric ratings, a structural equation model was fit to the self-ratings of female twins’ lifetime history of depressive symptoms and to the ratings of them by their cotwins, obtained through a self-report questionnaire. Our model allows the putative effects of questionnaire format to be estimated and distinguished from bias specifically related to the phenotype of the informant. Further, by including both these effects a ‘purer’ estimate of the genetic and environmental influences on depressive symptoms free from the effects of these two sources of error can be obtained.

METHODS

Subjects and measurements

The subjects included 463 monozygotic and 308 dizygotic Caucasian female twin pairs from the Virginia twin registry who had participated in a large population study of anxiety and depression. Extensive information on this sample has been previously reported (Silberg et al., 1990). Twins were asked to complete six questions regarding their lifetime history of depressive symptoms (their self-rating) and the lifetime history of depressive symptoms of their cotwin (rating of other). These questions were selected from the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) for diagnosing Major Depression (American Psychiatric Association, 1987). Specifically, the twins were asked to indicate ‘no’ (0) or ‘yes’ (1) to the following questions about themselves and their cotwin: ‘Had there ever been a period in your life lasting two weeks or more, when you/your cotwin:
1. felt depressed, blue, down in the dumps most of the time,
2. had trouble with poor appetite or overeating,
3. were troubled by low energy or fatigue,
4. really, really got down on yourself/herself and felt worthless,
5. had problems with poor concentration or difficulty making decisions,
6. had at least three of the problems listed above all at the same time’ (see Table 1).

Data analysis

Genetic and environmental components of variance obtained using the family history method versus the family study method. For exploratory purposes, we initially considered the degree to which the twin correlations and resulting estimates of the genetic and environmental influences on depressive symptoms may differ depending on whether data is collected via the family study method or the family history method. Because of its presumed clinical significance, we considered item 6, ‘had at least three of the problems listed above all at the time’ for this analysis.

Estimates of heritability, as well as estimates of the influence of the shared and non-shared environment, were obtained by fitting simple structural
Table 1. For each of the following questions, please circle the YES or the NO as appropriate. Answer all questions about you first, and then answer the questions about your twin. Once again, we are interested in how well you know your twin, so please do not consult with your twin.

<table>
<thead>
<tr>
<th>Has there ever been a period in your life, lasting two weeks or more, when you/your twin:</th>
<th>You</th>
<th>Your twin</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) felt depressed, blue, down in the dumps most of the time?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(2) had trouble with poor appetite or overeating?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(3) were troubled by low energy or fatigue?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(4) really got down on yourself/herself and felt worthless?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(5) had problems with poor concentration or difficulty making decisions?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>(6) had at least three of the problems listed above all at the same time?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Equation models to (a) the twin correlation matrices of both twins’ self-ratings (the family study method) and (b) the correlation matrices of a twin’s self-rating and their ratings of their co-twins (the family history method).

Heath et al. (1989) have detailed how structural equation modelling (employed in the software package LISREL7 (Joreskog and Sorborn, 1988)) may be applied to twin data to decompose the total variance of a particular trait into its genetic and environmental components. Specifically, it is the genetic and environmental relationships between monozygotic (MZ) and dizygotic (DZ) twins that allow us to estimate the magnitude of (1) additive genetic factors, those genetic effects that result from the additive effect of alleles at several loci, (2) shared environmental influences, those aspects of the environment that both twins share, such as parenting effects, socioeconomic status, and similar treatment by other people, and (3) non-shared environmental factors that affect only one member of the twin pair, such as physical illness, life events, short-term fluctuations in behaviour, and measurement error. Quantitative genetic theory predicts an additive genetic correlation of unity in MZ twins and 0.5 in DZ twins. By definition, the correlation between common environmental influences (i.e., family effects which are shared by both members of the twin pair) is unity for both zygoisty groups, while the correlation between specific environmental effects is zero.

Since twins share the same age, this variable could inflate the estimate of the shared environment if there is a relationship between age and the report of the lifetime prevalence of depressive symptoms. To determine the potential extent of age effects in the analysis of item responses, the statistical significance of the correlation between the twins’ age and the six items of the questionnaire was considered. Since the correlations between the age variable and the depressive items were both statistically non-significant (in both p > 0.20), age correction procedures were considered unnecessary for subsequent analyses of the twin data.

An ‘ACE’ model that estimates additive genetic (A), shared (C), and non-shared environmental (E) influences was fit to the correlation matrices of ratings of depressive symptoms by the method of weighted least squares. Because the twins’ ratings are discontinuous, the polychoric correlation was used, a statistic which assumes that the observed discontinuous ratings result when thresholds are superimposed upon an underlying normal distribution of liability (Olsson, 1979). The models were fit by the method of weighted least squares and evaluated using chi² statistics. To determine whether there is significant heterogeneity in the genetic and environmental estimates under the two methods, a combined model that constrained the self-ratings and the ratings of the cotwin to be equal was compared to the individual models using a likelihood ratio chi² test.

Modelling rater bias. After comparing the parameter estimates obtained using the family study method versus the family history method, a more comprehensive structural equation model was fit to the twin data, one which specifies hypothesized causal influences on both the twins’ self-ratings and their ratings of their co-twins’ depressive symptoms. This multivariate model not only allows
us to use information from both twins and from the
two different types of ratings (i.e., self ratings and
ratings of the cotwin) to obtain heritability
estimates of a latent ‘depressive’ phenotype, but it
also allows us to estimate the magnitude of rater
bias in psychiatric ratings and to determine the
extent to which individual rating styles and
questionnaire format may inflate perceived similari-
ity among relatives. As a result, the model is able to
provide an estimate of the genetic and environ-
mental influences on the latent rating of depressive
symptoms, free from these two bias effects. An
illustration of this proposed multivariate model for
rater bias is shown in Fig. 1. A modified version
of this model has been used previously to study bias in
the parent–offspring ratings of personality (Neale
BIAS IN RATINGS OF DEPRESSIVE SYSTEMS

and Stevenson, 1989) and twins' ratings of parent's educational attainment (Heath et al., 1985).

The rater bias model includes latent additive genetic factors (A), environmental factors shared by both members of a twin pair (C), and non-shared environmental factors specific to the individual (E). On two measured items of depressive symptoms for each twin, obtained directly (self-report) and indirectly (rating of cotwin). The selection of these two items, which is required for identification of the model, is detailed subsequently.

The latent genetic and environmental factors influence a second level of latent variables representing the 'true' depressive phenotype of the twins \(LP_1\) and \(LP_2\), which then influence the self-ratings of both twins on the two depression items \(S_{11}, S_{12}, S_{21}\), and \(S_{22}\). Partial regression coefficients, which represent the contribution of the path to the correlation between sets of latent variables and the latent and observed variables (Wright, 1968), include: (1) 'h', the path from additive genetic factors to the twins' latent phenotype, (2) 'c', the path from the shared environmental factors to the latent phenotype, (3) 'e', the path from the non-shared environmental factors to the latent phenotype, and (4) 's', the path from the latent phenotype to the twins' self-ratings.

The ratings of the cotwins' depressive symptoms \(R_{11}, R_{12}, R_{21},\) and \(R_{22}\) are influenced by latent rating variables \(LR_1\) and \(LR_2\), which are themselves influenced by the latent phenotype of the cotwin who is being rated (via an accuracy path \(a\)), the phenotype of the twin who is doing the rating (via the bias path \(b\)), and a residual error term \(R_e\). Additional influences on the observed ratings of depressive symptoms are error effects \(C_{14}\) and \(C_{25}\) which may be correlated \(r_e\) for self-ratings and ratings of cotwins on the same item.

To identify the rater bias model, the partial regression coefficient 'c' was fixed to unity. Since fitting to correlation matrices does not take account of the difference in variance between MZ and DZ twins predicted by the rater bias model, a correction factor (cf) was introduced (details available upon request).

As in the simple univariate analysis, item 6 of the depression questionnaire was used to fit the rater bias model. To choose the second item, a factor analysis was performed on the first five depressive items, and the item with the highest loading on the one common factor was selected.

The model was fit by the method of weighted least squares, using the asymptotic covariance matrices of the polychoric correlations as weights. Likelihood ratio tests were used to determine whether bias effects and correlated error could be removed from the model, without resulting in a significant deterioration in fit.

RESULTS

Table 2 presents the twin correlations for item 6 using the self-ratings and ratings of the cotwin (the family history method), and the self-ratings only (representing the family study method). This table shows that the correlation between twins for depressive symptoms is substantially higher when using the twin's self-ratings and their ratings of their cotwin \(0.724\) and \(0.858\) for MZ; \(0.572\) and \(0.562\) for DZ), as compared to using the self-ratings of both twins \(0.458\) for MZ; \(0.216\) for DZ.

Table 3 presents the results from fitting structural equation models to the correlations between the twins' self-ratings of depressive symptoms alone (family study method), between the twins' ratings of their cotwins, and the correlations between the twin's self-rating and rating of the cotwin (family history method). A comparison of these methods show that whereas the common environmental variance component is estimated as zero as a result of using the family study approach, it is significantly higher (0.31) when the family history method is used \(\chi^2 = 39.62, p < 0.001\).

There is no significant difference between the estimates as a result of using the twins' self-ratings as compared to using the twins' ratings of their cotwins (twin 1's rating of twin 2 and twin 2's rating for twin 1), \(\chi^2 = 0.25, p = 0.98\).

The first item of the questionnaire, 'felt depressed, blue, down in the dumps most of the time' had the highest loading on the single common factor derived from a factor analysis of the first five depression items of the questionnaire. This item and item 6 were included as indices of the twin's ratings of depressive symptoms for fitting the more comprehensive rater bias model.

The fit of the full bias model for the two depression items was excellent \(\chi^2 = 57.98, p = 0.55\). A reduced rater bias model with 3 fewer parameters (ie constraining 's', 'c', and 'r' to be equal for the two items) did not result in a significant worsening of fit compared to the full model (the likelihood ratio \(\chi^2 = 0.98, p = 0.81\)). Since a simpler model with fewer parameters could
Table 2. A comparison of twin correlations derived using the family history method and the family study method. \(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Twin1self</th>
<th>Twin2byTwin1</th>
<th>Twin2self</th>
<th>Twin1byTwin2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twin1self</td>
<td>—</td>
<td>0.572(^a)</td>
<td>0.216(^a)</td>
<td>0.367</td>
</tr>
<tr>
<td>Twin2byTwin1</td>
<td>0.724(^a)</td>
<td>—</td>
<td>0.496</td>
<td>0.183</td>
</tr>
<tr>
<td>Twin2self</td>
<td>0.458(^b)</td>
<td>0.392</td>
<td>—</td>
<td>0.562(^b)</td>
</tr>
<tr>
<td>Twin1byTwin2</td>
<td>0.578</td>
<td>0.419</td>
<td>0.858(^a)</td>
<td>—</td>
</tr>
</tbody>
</table>

\(^a\)MZ correlations presented in lower triangle.  
DZ correlations presented in upper triangle.  
\(^b\)Family history method.  
\(^c\)Family study method.

Table 3. Estimates of phenotypic variance components using twins’ rating of self, rating of cotwin, and combined self-rating and rating of cotwin. \(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Self-rating</th>
<th>Rating of cotwin</th>
<th>Self-rating with rating of cotwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes</td>
<td>0.45</td>
<td>0.41</td>
<td>0.50</td>
</tr>
<tr>
<td>Shared environment</td>
<td>0.00</td>
<td>0.00(^a)</td>
<td>0.31</td>
</tr>
<tr>
<td>Non-shared environment</td>
<td>0.55</td>
<td>0.59</td>
<td>0.19</td>
</tr>
</tbody>
</table>

\(^a\) \(\chi^2\) test of heterogeneity between family study method (Self-Ratings) versus family history method (Self Rating with Cotwin’s Rating) = 39.62 \(p < 0.001\).  
\(^b\)Estimate is constrained at a lower bound of zero.

sufficiently account for the twin data, the more parsimonious model was used. However, dropping either ‘b,’ the bias parameter, or ‘c,’ the correlated item error parameter did result in a significant worsening of fit (\(\chi^2\) = 113.90 and 84.61, respectively, \(p < 0.001\)).

The components of phenotypic variance under the reduced rater bias model for the twins’ latent ‘depressive’ phenotype, the twins’ self-ratings, and twins’ ratings of their cotwins (for item 1 and item 6) are presented in Table 4. After correcting for the various sources of bias, approximately 29% of the total variation underlying the twins’ latent phenotype for the rating of lifetime prevalence of depressive symptoms can be explained by additive genetic factors, 20% to those environmental experiences shared by both members of the twins pair, and the remaining 51% to those environmental effects specific to the individual. The estimate of the bias parameter (b) is moderately large and positive, and accounts for approximately 20% of the total variation in the twins’ ratings of their cotwins. The accuracy parameter (a) is smaller, and explains 7% of the variance of twins’ ratings, random error (r) 42%, and the correlated error between the twins’ self-rating and rating of her cotwin (c) explains 18%. The remaining variance, 0.13% (mz) and 0.07% (dz), can be accounted for by the covariance between the twins. Because of the inclusion of item error, the sources of variation for the twins’ self-ratings are somewhat lower than those for the latent phenotype. However, the similarity of the variance components underlying the twins’ latent depressive phenotype as compared to the twins’ self-ratings is noteworthy.

**DISCUSSION**

By fitting a structural equation model to female twins’ self-ratings of lifetime history of depressive symptoms and the ratings of them by their cotwins, we were able to examine whether ratings of depressive symptoms are systematically affected by characteristics of the rater. Parameter estimates under the model indicate that at least one fifth of the variance in the twins’ ratings of their cotwins’
Table 4. Variance components for twins latent 'depressive' phenotype, self rating, and rating of cotwin.

<table>
<thead>
<tr>
<th></th>
<th>Latent phenotype</th>
<th>Self rating</th>
<th>Rating of cotwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes</td>
<td>0.29</td>
<td>0.24</td>
<td>-</td>
</tr>
<tr>
<td>Shared environment</td>
<td>0.20</td>
<td>0.17</td>
<td>-</td>
</tr>
<tr>
<td>Non-shared environment</td>
<td>0.51</td>
<td>0.42</td>
<td>-</td>
</tr>
<tr>
<td>Correlated error</td>
<td>-</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Random error</td>
<td>-</td>
<td>-</td>
<td>0.42</td>
</tr>
<tr>
<td>Bias</td>
<td>-</td>
<td>-</td>
<td>0.20</td>
</tr>
<tr>
<td>Accuracy</td>
<td>-</td>
<td>-</td>
<td>0.07</td>
</tr>
<tr>
<td>Twin covariance</td>
<td>-</td>
<td>-</td>
<td>0.13_{[MZ]}, 0.07_{[DZ]}</td>
</tr>
</tbody>
</table>

history of depression, can be accounted for by rater bias, lending support to the hypothesis that twins tend to rate cotwins according to their own history of depressive symptoms. Those twins who report having experienced few or no depressive symptoms rate their cotwins as having fewer lifetime symptoms than those cotwins rate themselves, whereas those twins with a history of depressive symptoms perceive their cotwin as having more depressive symptoms than the cotwin perceives.

An important aspect of the present analysis was also to distinguish the similarity among twins due to the rater's projection of depressive symptoms (rater bias) versus the similarity due to the tendency of twins to perceive themselves similar to their cotwin in a variety of ways and therefore rate their cotwin like themselves on most any characteristic. Since the latter similarity can be exacerbated by a questionnaire format that facilitates the replication of responses of the self-rating and the rating of other by asking the informant to rate her subject on the same occasion and/or the same page of a questionnaire, we allowed for the item error to be correlated between the twins' self ratings and their ratings of their cotwin in our analysis. By doing so, we were able to estimate that proportion of the total variance in the twins' ratings attributable to questionnaire format and a general misperception of familial similarity and distinguish it from the bias related to informant's history of depressive symptoms. The moderately high correlation found between the error in twins' ratings (0.27) underscores the importance of question format when asking relatives for diagnostic information regarding the psychiatric status of family members. We attribute a large portion of the correlated error to the placement of the twins ratings in the questionnaire. Our results suggest that this type of error could be minimized by separating the two ratings in time and in place.

The tendency of raters to project their own experience onto the subjects they are rating has important implications for the collection of data by those methods that rely upon ratings from relatives for diagnosing psychiatric illness in family members, such as the family history method. We predict that these ratings will be biased. The estimates obtained when relatives' ratings were used suggests that overall, the estimate of familial transmission of depression may be spuriously increased using the family history approach as a result of the number of false negatives (those people diagnosed as well, when they are not) derived from informants who have no history of depressive symptoms.

The advantage of the present model is that it can correct for various sources of bias when estimating the genetic and environmental influences underlying a trait. In its present application, the model was fit to data in which the self-ratings and the informant's ratings of depressive symptoms for all individuals were available. Since the model can correct for bias inherent in informants' ratings, information regarding the diagnostic status of individuals who are not available for interview can also be incorporated in this type of analysis, using ratings of these individuals by other family members. However, the model does require that at least a portion of the sample being studied include individuals where both sets of ratings (that is, self-ratings and ratings of relatives) are available.

The twin design has been used in the present analysis because of its ability to provide reliable estimates of the genetic and environmental parameters underlying ratings of depressive symptoms. However, the rater bias model can be extended to estimate the genetic and environmental influences on psychiatric symptoms using ratings from other sets of informants, such as siblings and parents and offspring, or others for whom the genetic and environmental relationship is known.

Besides its obvious relevance to data collection, the finding of bias in the ratings of depressive symptoms is also relevant to the psychological literature on depression and perceptual processing.
Beck (1967) has postulated that depression is a direct consequence of negative cognitions about oneself, one’s environment, and the future. His theory has been supported in a number of studies investigating the influence of experimentally induced depression on the perception of negative information regarding the self (Blaney, 1986; Bower, 1981) and depressives’ recall of negative parental behaviour (Lewinsohn and Rosenbaum, 1987). Opposite to what Beck’s theory would predict, Lewinsohn and Rosenbaum demonstrated that the negative perception of parental behaviour was a concomitant rather than a precipitant to depression. Those individuals who were currently depressed and those who were soon to be depressed remembered their parents more negatively than those individuals who were not currently depressed, but who had a history of depression.

The finding that those twins having experienced depressive symptoms tend to rate their relatives as having experienced more depressive symptoms is consistent with Beck’s theory of negative recall of the environment by those individuals who are prone to experience depressive symptoms. However, we did not find negative cognitions to be a consequence of the current depressed mood state, as hypothesized by Lewinsohn. Rather, the inclusion of the current mood state of the twin, as measured by the Center for Epidemiological Studies Depression Scale (CES-D) (Comstock and Helsing, 1976; Radloff, 1977), in the rater bias model had little impact on the twin’s perception of the cotwin’s lifetime history of depressive symptoms.

Bias has also been found in several other studies investigating the genetic and environmental effects on ratings of parental educational attainment (Heath et al., 1985), perception of body size of family members (Meyer, 1987), the perception of social attitudes among identical and non-identical twins (Eaves and Last, 1980), and in spouses’ ratings of the temperament of their offspring (Neale and Stevenson, 1989). The overestimation of phenotypic similarity in these studies underscores the potentially deleterious effect rater bias can have on obtaining reliable data across a variety of different phenotypes.

The importance of additive genetic and shared and non-shared environmental factors in the report of lifetime prevalence of depressive symptoms was also demonstrated in the present analysis. The predominance of non-shared environmental factors (those environmental effects experienced by only one member of the twin pair) is similar to that found by both Kendler et al. (1987) and Silberg et al. (1990) in a multivariate analysis of depressive symptoms. The estimate of 29% genetic variance for the latent phenotype for depressive symptoms is consistent with the 27% reported by Kendler et al. (1987) for the first common genetic factor influencing the covariation among items representing symptoms of anxiety and depression from the Delusions–Symptoms–States Inventory (Bedford et al., 1976). The finding of a strong shared environmental component (those environmental events common to both twins) in the report of twins’ lifetime history of depressive symptoms is surprising, however, in light of the results of previous studies. Neither Kendler et al. (1986; 1987) nor Silberg et al. (1990) found strong evidence for the role of the common environment, such as parenting style or other shared familial experiences, in the report of depressive symptoms in adulthood. It is possible that this discrepancy is due to the context in which the depressive symptoms are rated. The scales analysed by Kendler et al. (1986; 1987), and Silbert et al. (1990) instruct twins to rate depressive symptoms over a short time period, whereas the present study analyses the lifetime history of symptoms of the twins. One possibility is that these lifetime assessments are more sensitive to the effect of early shared environmental experiences than ones that address symptoms only in the recent history of the twin. However, given the incomparability between the items of the scales of these studies, it is impossible to test this assumption directly at this time.

The discrepancy between estimates of common environment in the simple ACE model in comparison to the rater bias model also deserves comment. The difference between the two models is not surprising given the greater complexity of the rater bias model, which analyses two depressive symptoms concurrently and utilizes information from twins’ ratings of both themselves and their cotwin.

We do not know the applicability of the present study to clinical depression. We are presently collecting diagnostic data on twins and their families that will be available for analysing ratings of clinical depression and anxiety.

Given the widespread use of the family history method for analysing familial resemblance for psychiatric illness, the suggested method for controlling for various sources of bias in the estimation of the genetic and environmental components of
variance should be useful to future genetic studies of psychiatric disorders.

ACKNOWLEDGEMENTS

This work was made possible through National Institute of Mental Health Grants MH-40828, MH-45268, and GM-30250.

REFERENCES

American Psychiatric Association (1987) Diagnostic and Statistical Manual of Mental Disorders (3rd edn, revised), Washington DC.