The Genetic Epidemiology of Self-Esteem

MARC-ANDRÉ ROY, MICHAEL C. NEALE and KENNETH S. KENDLER

Background. Previous studies on self-esteem have focused exclusively on its psychosocial determinants. The goal of the present study is to clarify genetic v. environmental determinants of self-esteem.

Method. Participants were Caucasian women sampled from the Virginia Twin Register: 363 pairs of MZ and 238 pairs of DZ twins were available from the first wave of the study, and 430 pairs of MZ and 308 pairs of DZ twins from the second. Self-esteem was assessed with the Rosenberg’s Self-Esteem Scale.

Results. Using univariate twin analyses of self-esteem and a repeated measurement twin model, we found that self-esteem is a moderately heritable trait (heritability = 52% in the repeated measurement model); environmental influences are also very important, and are probably mostly not shared by members of a twin pair.

Conclusions. Aetiological models of self-esteem which examine only psychosocial factors are incomplete; genetic factors need to be integrated.

Self-esteem is a concept often used in both clinical and research contexts (Robson, 1988). Indeed, a literature search in a computerised bibliography covering the last 25 years (Psych-Info) identified over 20 000 entries under this heading. Frequently cited consequences of poor self-esteem include interpersonal dependency (Hirschfeld et al., 1976), delinquent behaviour (Gold, 1982), substance abuse (Brehm & Back, 1968), eating disorders (Yates, 1989), anxiety disorders (Ingham et al., 1986) and depression (Brown & Harris, 1978; Beck et al., 1980).

Given this interest in the role of self-esteem, it is not surprising that much research has been devoted to the psychosocial determinants of individual differences. According to the same body of literature, the determinants of poor self-esteem would include unemployment (Warr & Jackson, 1983), lack of social support (Brown & Harris, 1978), early parental loss (Brown & Harris, 1978), and rearing environment characterised by (i) parents' lack of acceptance of the children, (ii) poor ability to set and enforce limits (Coopersmith, 1967) and (iii) parental conflict (Watkins, 1976). However, research on self-esteem is complicated by important differences between scales (Bingham, 1983; Robson, 1988): some scales (e.g. Rosenberg’s, 1965) reflect a global self-approval, whereas others (e.g. Brown et al., 1990) assess competency in specific roles. A second difference is whether the scales include items related to both poor self-appraisal (i.e. endorsement of items reflecting poor self-esteem) and lack of good self-appraisal (i.e. failure to endorse items reflecting good self-esteem). These are thought to represent two distinct dimensions; this is supported by factor analytic studies (Bagley & Evan-Wong, 1975; Dobson et al., 1979) and by studies finding different correlates for each of these dimensions (Bagley & Evan-Wong, 1975). These and other discrepancies between scales probably partly explain the poor correlation observed between them (Kahle, 1976).

In contrast to the abundance of studies on psychosocial determinants, we are aware of no published family, twin or adoption study of the familial/genetic determinants of self-esteem. The importance of these factors is suggested by twin and adoption studies of various personality traits, which show some degree of genetic control for most traits (McGuffin & Thapar, 1992). Indeed, twin studies seek to explain the variations among individuals by: (i) additive genetic effects (A); (ii) environmental influences common to both twins (C, or shared environment); and (iii) environmental influences specific to each twin (E, or non-shared environment). Thus, C would reflect factors shared by both members of a twin pair, such as early parental loss and rearing environment, while E would reflect factors not shared by both twins, such as stressful life events of adulthood (such as unemployment) and lack of social support. Twin studies could, therefore, also shed light on the environmental influences on self-esteem.

The goal of this paper is to quantify the relative importance of genetic and environmental influences on self-esteem, using cross-sectional and repeated-measurement twin models.

Method

Methods have already been described in more detail elsewhere (Kendler et al., 1992). Briefly, subjects
were Caucasian female twin pairs selected from the population-based Virginia Twin Register, which is based on a systematic review of the birth records in the state of Virginia. The present results are based on the first two waves of the study. In wave I, questionnaires covering self-esteem as well as psychiatric symptoms, personality traits and risk factors for psychiatric disorders were mailed to these subjects; the response rate was 64%. Two different versions of the questionnaire were sent to distinct subgroups of the participants; one of these versions did not include self-esteem. We were left with 363 complete monozygotic (MZ) pairs and 238 dizygotic (DZ) pairs at time 1. In wave II, which took place on average 16 months later, the women who participated in wave I were asked to come to a face-to-face personal interview, to which 92% (2163 individuals) responded; but since our analyses are based only on questionnaire data, we will not here describe this interview. The subjects were also asked to complete a questionnaire including a self-esteem scale; 85% (1842, for 430 complete MZ and 308 complete DZ pairs) of the subjects who were interviewed returned this second questionnaire.

Zygosity was assigned by an algorithm that included, for most twins, a review of photographs and the responses to questions about physical similarity and frequency of confusion as children. When necessary, DNA testing was performed as well.

Self-esteem was measured using the full Rosenberg Self-Esteem Scale (RSES, Table 1) (Rosenberg, 1965), in a self-administered questionnaire format. This scale was chosen because it had been used extensively in previous studies and shown to be reliable and valid (see Wylie, 1989, for a comprehensive review). The subjects were asked whether they strongly agreed, agreed, disagreed or strongly disagreed with the 10 statements in Table 1, which were interspersed with ones from other scales. The scores on the five items phrased positively were inverted, so that a high score consistently reflected low self-esteem.

To test the Equal Environment Assumption (EEA) (i.e. that MZ are not more correlated than DZ for exposure to aetiologically significant environmental exposures), we assessed the degree of environmental similarity of the twins in childhood and their frequency of contact in adulthood (Kendler et al., 1993d).

Statistical methods
Factor analysis of self-esteem
Given some discrepancies among previous factor-analytic studies of the RSES (Kaplan & Pokorny, 1969; Hensley & Roberts, 1976; Dobson et al., 1979), we repeated these analyses on the present sample. Orthogonal (Varimax) and oblique (Promax) factor analyses were performed using the principal component method in the Statistical Analysis System (SAS Institute, 1990) on twins who had answered all items on the RSES. Factors were considered meaningful if their eigenvalue was greater than 1 and if they appeared on the vertical part of the scree plot.

Genetic analyses
We computed variance–covariance matrices for twin 1’s and twin 2’s self-esteem for MZ and DZ twins, using PRELIS (Jöreskog & Sörbom, 1988). Genetic models were fitted on these matrices using maximum likelihood methods as implemented in MX (Neale, 1991).

Univariate models
Figure 1 shows a path diagram of a simple univariate model. In the full model, both additive genes (A) and shared environmental factors (C) (Neale & Cardon, 1992) cause twins to be more similar to each other than unrelated individuals. However, A and C predict different patterns of resemblance. First, since MZ twins share all their genes and DZ twins, on average, only half, A will cause the intrapair correlation in MZ twins to be twice as large as that of DZ twins. Second, since we are assuming that MZ and DZ twins share their environment to the same extent, C will cause the correlation in MZ and DZ.

![Fig. 1 Path diagram of a simple univariate model of self-esteem. In the full model, resemblance between twins is assumed to result from two sets of latent (i.e. unobserved) factors: (i) additive genes (A), which causes the correlation between twins to be twice that of DZ, since MZ share all their genes and DZ only half; (ii) environmental factors shared by both twins (C or shared environment), which tends to make the correlation between MZ and DZ similar. Conversely, this model postis a third factor, environmental influences not shared by the twins (E or individual specific environment). These latent factors (i.e. A, C and E) influence the measured variables through the paths a, c and e respectively. The variance of self-esteem is partitioned in a proportion a', c' and e' reflecting each of these paths.](image-url)
The genetic epidemiology of self-esteem
twins to be equal. This univariate model posits a third factor, environmental influences not shared by the twins (E or individual specific environment). Since E is not shared by the twins, it will be reflected by the correlation in MZ being less than 1, i.e. it will tend to make members of a twin pair no more similar than pairs of unrelated subjects. In cross-sectional analyses, E also reflects measurement error and short-term fluctuations. A, C and E are termed latent factors since they are unobserved. They are derived statistically, based on the patterns of resemblance in MZ v. DZ twins. A, C and E influence the measured variables through the paths a, c and e respectively. Since the path coefficients (a, c and e) were standardised, they can be interpreted as correlation coefficients, and the variance of self-esteem was partitioned into proportions a^2, c^2 and e^2, reflecting each of these paths.

Using MX, one specifies the genetic model to be tested (e.g. ACE, AE, CE or E). MX will then go through successive iterations until it has reached values of the parameters (e.g. a, c or e) that maximise the fit of the model tested with the data. Then, the fit of that particular model is assessed by a goodness-of-fit χ^2, a smaller χ^2 (and a larger P value) meaning better fit. The choice of the preferred model was based on the Akaike criterion (AIC) (Akaike, 1987). Since the AIC is obtained through the expression χ^2 - (2\*number of degrees of freedom), it reflects both the goodness of fit and the parsimony of the model, and models with smaller (i.e. more negative) AIC are preferred.

We first fitted a full ACE model. Then, the effect of dropping one of the parameters is statistically tested by computing the difference in χ^2 between the models. The AE model postulates that all familial aggregation results from additive genetic effect (A), while the CE model postulates that all familial aggregation results from shared environmental influences (C). The E model postulates no familial transmission. We also fitted ADE models, but since the effect of D (dominance) was never significant, these models will not be further discussed.

Repeated measurement model

In cross-sectional univariate models, E is confounded with random error and with short-term variation (Cronbach, 1970; Kendler et al, 1993c). However, using repeated measurements may allow us to examine the genetic architecture of the component of self-esteem which is stable over two occasions, distinguishing random error and short-term fluctuations from longer term environmental influences (Fig. 2, and Kendler et al, 1993c). The repeated measurement model assumes that there is a 'true' (or stable) level of self-esteem, which is imperfectly reflected by any single measurement. This latent (or stable) level of self-esteem (P in the figure) is modelled as in the standard twin design, using the latent factors A, C and E, which influence the stable component through paths a, c and e. This latent phenotype then influences the observed variables through path λ. Therefore, λ reflects the proportion of the variance in self-esteem that is explained by the stable component. Conversely, the path k reflects either measurement error or short-term fluctuations in self-esteem. Thus, k^2 is the proportion of the variance accounted for by unreliability or short-term fluctuations. The model is constrained so that λ^2 + k^2 = 1. We also tested for correlation of error between twins by adding a path from twin 1 and twin 2 latent error factors at time 1 and at time 2. Since these paths were not significant, they will not be discussed further.

Test of the Equal Environment Assumption

The method we used to test the EEA has been described previously (Kendler et al, 1993d). We used linear regression, with independent variables
(i) zygosity and (ii) the measures of similarity of environment or frequency of contact, and the absolute value of the within-pair difference of self-esteem as dependent variable. Thus this method seeks to examine whether similarity of the environment correlates with the phenotypic similarity between twins, controlling for zygosity. A violation of the EEA would be detected by similarity of environmental exposure significantly predicting similarity of self-esteem.

Results

Test of biases and EEA

Using Pearson correlation coefficients, we found no correlation between age and self-esteem at Time 1 or Time 2. Using t-tests, we found no significant difference on the mean of self-esteem between either MZ and DZ twins, or members of complete v. incomplete pairs; since members of twin pairs are correlated for self-esteem, the participating member of an incomplete pair provides indirect information on the non-participating member. This suggests that the subjects who refuse to participate do not systematically differ from those who agree. Moreover, neither frequency of contact or similarity of environment significantly predicted similarity in self-esteem, which suggests that there was no major violation of the EEA.

Factor analysis and scale construction

Table 1 shows the loadings on the two meaningful factors obtained from a Varimax orthogonal rotation.

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I On the whole, I am satisfied with myself*</td>
<td>0.66‡</td>
<td>0.30</td>
</tr>
<tr>
<td>II At times, I think I am no good at all</td>
<td>0.77</td>
<td>0.09</td>
</tr>
<tr>
<td>III I feel I have a number of good qualities*</td>
<td>0.23</td>
<td>0.72</td>
</tr>
<tr>
<td>IV I am able to do things as well as most other people*</td>
<td>0.07</td>
<td>0.77</td>
</tr>
<tr>
<td>V I feel that I do not have much to be proud of</td>
<td>0.54</td>
<td>0.32</td>
</tr>
<tr>
<td>VI I certainly feel useless at times</td>
<td>0.76</td>
<td>0.06</td>
</tr>
<tr>
<td>VII I feel that I am a parson of worth, at least equal with others*</td>
<td>0.26</td>
<td>0.71</td>
</tr>
<tr>
<td>VIII I wish I could have more respect for myself</td>
<td>0.76</td>
<td>0.13</td>
</tr>
<tr>
<td>IX All in all, I am inclined to think that I am a failure</td>
<td>0.65</td>
<td>0.31</td>
</tr>
<tr>
<td>X I take a positive attitude toward myself</td>
<td>0.69</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*The score on items phrased positively is reversed.
†Italics indicate items which were used in later analyses.

Using the Varimax rotation, factor 1 explained 35% of the variance, and factor 2 explained 20% of the variance. The scree plot and an oblique Promax rotation also suggested that these were two meaningful factors. The first factor had highest loadings on items 1, 2, 5, 6, 8, 9 and 10; all but two of these items are negatively phrased. Conversely, the second factor had highest loadings on items 3, 4 and 7, which are all worded positively. Our analyses were based on a summation of the seven items loading mainly on factor 1. Factor 2 was not analysed further, since it would suffer from substantial measurement error, because it is based on only three items.

Univariate analyses of self-esteem

Table 2 shows the results for univariate analyses of self-esteem at Time 1 and at Time 2. In both instances, C could be eliminated without any deterioration of fit ($\chi^2 = 0.04$, 1 d.f., $P = 0.84$ at Time 1, $\chi^2 = 0.00$, 1 d.f., $P = 1.00$ at Time 2), which means that a model of genetic transmission (AE model) explained the data as well as a model of mixed genetic and environmental transmission (ACE). However, a model of purely environmental transmission (CD) caused a significant deterioration of fit ($\chi^2 = 7.49$, 1 d.f., $P = 0.006$ at Time 1, and $\chi^2 = 14.20$, 1 d.f., $P = 0.007$ at Time 2). Therefore the AE model was preferred in both occasions, and yielded a heritability estimate of 0.40 at Time 1 and 0.36 at Time 2.

Repeated measurements of self-esteem

The test–retest reliability of self-esteem was 0.75 for MZ and 0.74 for DZ (Z test for comparing the correlation coefficients = 0.52, $P < 0.25$). The ACE offered a very good fit ($\chi^2 = 12.69$, 14 d.f., $P = 0.55$). We dropped C (AE model, third column) without any deterioration in fit ($\chi^2$ v. ACE = 0.00, 1 d.f., $P = 1.00$), but the elimination of A (CE model) caused a significant reduction of fit ($\chi^2 = 8.14$, 1 d.f., $P = 0.004$). Thus the AE model was the preferred one and yielded a heritability estimate of 33%. Its parameters are shown in Table 3. The reliability path ($\lambda$) was high on both occasions, accounting for approximately 75% of the variance.
Table 3
Best-fitting repeated-measures model of self-esteem

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Fit of the model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a^2$</td>
<td></td>
</tr>
<tr>
<td>$e^2$</td>
<td></td>
</tr>
<tr>
<td>$\lambda^2$ time 1</td>
<td>$12.69 (15 \text{ d.f.})$</td>
</tr>
<tr>
<td>$\lambda^2$ time 2</td>
<td>$0.72$</td>
</tr>
<tr>
<td>$k^2$ time 1</td>
<td></td>
</tr>
<tr>
<td>$k^2$ time 2</td>
<td></td>
</tr>
</tbody>
</table>

MZ time 1-time 2 correlation for self-esteem = 0.75.
DZ time 1-time 2 correlation for self-esteem = 0.74.
$e^2$: proportion of variance explained by additive genetic effects.
$a^2$: proportion of variance explained by environmental influences not shared.

Discussion

The main findings are that self-esteem, as measured by the RSES, is a relatively stable trait (test–retest reliability = 0.75), with moderate heritability (52% in the repeated measurement model which corrects for measurement error); and that environmental influences are also very important, and most of them are probably not shared by members of a twin pair.

Conceptualising self-esteem as a substantially heritable trait may appear counter-intuitive. First, genetic influences have been found for many different personality traits or attitudes, such as extraversion and neuroticism (Eaves et al., 1989), coping strategies (Kendler et al., 1991) and suspiciousness (Kendler et al., 1987). Thus self-esteem would only be one more trait in that list. A second argument would be that self-esteem is correlated with other traits which have been found to be heritable, i.e. that a trait correlated with a heritable trait is likely (but not certain!) to be heritable. Such traits include neuroticism (Bagley & Evan-Wong, 1975; Eaves et al., 1989), depressive symptoms (Kaplan & Pokorny, 1969; Silberg et al., 1990) and anxiety symptoms (Kaplan & Pokorny, 1969; Jardine et al., 1984; Kendler et al., 1986).

Since genes that directly influence self-esteem are very unlikely, the next logical question is: ‘what is inherited?’ A first possibility would be that a tendency toward chronic dysphoria is inherited, and that poor self-esteem arises as a consequence at the cognitive level, so that the heritability of self-esteem would reflect that of dysphoric mood. This is supported by finding that depressive symptoms, which could be thought of as reflecting such a dysphoric state, are significantly correlated with self-esteem and moderately heritable. Such a conceptualisation is reminiscent of Kraepelin’s depressive temperament (Kraepelin, 1921) or Kretschmer’s melancholic type (Kretschmer, 1950). These temperaments are said to be characterised by a ‘permanent gloomy emotional stress in all the experiences of life’ (Kraepelin, 1921) and they include poor self-esteem as a central feature.

A second explanation of the genetic influences on self-esteem would be that they reflect the impact on the environment of some temperamental feature. In such a model, the reaction of persons in the environment to one’s temperament would be a crucial factor in determining one’s self-esteem. For example, high neuroticism scores have been shown to predict poor social support (Kendler et al., 1993a). In turn, such poor social support could be ‘introjected’ by the individual and result in poor self-esteem. Since neuroticism has been shown moderately heritable (Eaves et al., 1989), the heritability of self-esteem could be mediated through its relationship with neuroticism. This could be an instance of what has been named “genetic control of exposure to the environment” (Kendler & Eaves, 1986). Such effects of genotype on the environment are supported by finding substantial heritability for features of the environment such as life events (Kendler et al., 1993b; McGuffin et al., 1988) and social support (Bergeman et al., 1990; see Plomin & Bergeman, 1991, for a review).

One way to examine the relationship between self-esteem and depression or neuroticism is through multivariate twin analyses, which allow us to examine the extent to which two traits share the same genetic, shared environmental or specific environmental determinants (Vogler, 1985). We have completed preliminary analyses (which will be fully reported in a future publication), which included DSM–III–R (American Psychiatric Association, 1987) diagnosis of depression, based on a structured interview which we (Kendler et al.) described in 1992, and a questionnaire-based measurement of neuroticism (described in Heath et al., 1992). These analyses revealed that each of the following genetic factors explained approximately one-third of the genetic variance in self-esteem: (i) a factor common to neuroticism, self-esteem and depression; (ii) a factor common to self-esteem and depression; and (iii) a factor specific to self-esteem. The genetic influences in self-esteem are, therefore, likely to be complex, and will need to be explored further using multivariate genetic models.
The present study is also informative about the nature of environmental influences on self-esteem. These are substantial; even correcting for unreliability of measurement, they accounted for roughly 50% of the total variance. This is consistent with the general finding in previous studies that environmental influences account for at least half of the variance of psychological or psychiatric traits. Second, environmental influences on self-esteem are mostly not shared by members of a twin pair. The predominance of E over C agrees with results of previous studies on other personality features (Plomin & Bergeman, 1987; Eaves et al., 1989). However, our conclusion about the absence of a significant effect of shared environment should take into account: (i) the limited power of twin studies to detect modest effects of C in the presence of substantial genetic influences (Martin et al., 1978); and (ii) the confounding of C with genetic dominance (D) effects in the twin studies (Grayson, 1989). (But the full model yielded an estimate of C which was null, which renders a substantial effect of C unlikely.)

Third, environmental factors exert a largely stable effect, since E explained a large part of the variance even in repeated measurement models. This argues against short-term fluctuations in the environment being able to explain a large part of the variance in self-esteem. Our results are, therefore, compatible with: (i) long-term effects of childhood environment not shared by twins (e.g. parents’ attitudes towards their children, to the extent that they do not have the same attitude towards each of them); (ii) enduring conditions of adult life (e.g. chronic unemployment, chronic marital discord); (iii) lasting effects of discrete events in adult life (e.g. marital separation).

Implications

These findings warrant some general comments. Whatever the mechanism of genetic transmission of self-esteem, the present findings argue strongly against self-esteem being exclusively the result of environmental influences. They show how important it is that psychological models of the development of self-esteem should integrate genetic/temperamental factors.

Good self-esteem may be one of the mechanisms of the genetic influences on resilience to stress. Previous investigations have shown that people with good self-esteem are less likely to develop psychiatric distress when faced with stressful events (Rutter, 1985). Such people adopt an optimistic stance when faced with adversity, and have a better sense of control over their life than individuals with low self-esteem. However, this study did not directly address that issue, and this suggestion should therefore be seen as speculative.

The importance of genetic determinants should not be misinterpreted as meaning that low self-esteem cannot be improved, even if our repeated-measurement model suggests that genetic influences are stable over time. The following facts suggest that environmental manipulations may affect self-esteem: (i) psychotherapeutic experiments have shown improvements of self-esteem in subjects with low self-esteem during psychotherapies specifically targeting that problem (Gauthier et al., 1983); (ii) self-esteem is only moderately heritable, and almost 50% of the variance was explained by environmental factors; (iii) even a trait as decisively genetic as phenylketonuria can be alleviated through environmental manipulations.

Limitations

Any interpretation of the present results should bear in mind the following limitations.

First, only Caucasian women were included in the present sample. Previous studies have shown gender and race to influence self-esteem (Bingham, 1983; Robson, 1988), so we should be careful before generalising our results to other populations.

Second, owing to the discrepancies between scales mentioned in the introduction, we do not know if our present findings would apply to other measures of self-esteem.

Third, it is possible that aspects of the environment other than the similarity of childhood environment and the frequency of contact violated the EEA. However, previous investigations of the EEA in behaviour genetics have found very little support for significant violations of the EEA (Kendler et al., 1993a), so that such a bias probably did not play a major role in the present study.

Fourth, even if we used a population sample, some selection bias may have occurred. We have partially addressed that issue by comparing subjects from complete twin pairs with subjects from incomplete twin pairs. Given the resemblance between members of a twin pair, the participating subject from a complete pair provides indirect information on the non-participating member. However, it is still possible that some non-participation bias may have occurred, and the potential effect of this on the parameter estimates could be very difficult to assess (Neale & Cardon, 1992).

Acknowledgements

This research was supported by grants MH-40828 and AA09095. The Virginia Twin Registry, established and maintained by W. Nance and L. Corey, is supported by grant NS-25630 from
References


ROY ET AL


M.-A. Roy, FRCP, Department of Psychiatry, Medical College of Virginia, Centre de Recherche Université Laval Robert-Giffard, Beaupre, Quebec; M. C. Neale, PhD, Department of Psychiatry, Medical College of Virginia; K. S. Kendler, MD, Departments of Psychiatry and Human Genetics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA, USA

Correspondence: Dr Roy, Centre de Recherche Université Laval Robert-Giffard, 2601 de la Canadiere, Beaupre, Province de Quebec, Canada G1J 2O3

(First received 14 February 1994, final revision 14 September 1994, accepted 7 October 1994)