Journal of Personality and Social Psychology
1992, Vol. 63, No. 1, 85–96

Copyright 1992 by the American Psychological Association, Inc.
0022-3514/92/$3.00

Evidence for Genetic Influences on Personality From Self-Reports and Informant Ratings

A. C. Heath
Department of Psychiatry, Department of Psychology, and Department of Genetics
Washington University

M. C. Neale
Department of Human Genetics
Virginia Commonwealth University

R. C. Kessler
Institute for Social Research
University of Michigan

L. J. Eaves and K. S. Kendler
Department of Human Genetics and
Department of Psychiatry
Virginia Commonwealth University

Self-report data on Extraversion (E) and Neuroticism (N), together with ratings by the co-twin, were obtained from a sample of 826 adult female twin pairs ascertained through a population-based twin register. Data were analyzed using a model that allowed for the contributions to personality ratings of the rater's personality (*rater bias*) as well as of the personality of the person being rated. For E, but not for N, significant rater bias was found, with extraverted respondents tending to underestimate, and introverted respondents tending to overestimate, the Extraversion of their co-twins. Good agreement between self-reports and ratings by the respondent's co-twin was found for both E and N. Substantial genetic influences were found for both personality traits, confirming findings from genetic studies of personality that have relied only on self-reports of respondents.

Findings from adoption studies (Loehlin, 1981; Loehlin, Willerman, & Horn, 1985; Scarr, Webber, Weinberg, & Wittig, 1981), from separated-twin studies (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Pedersen, Plomin, McClearn, & Friberg, 1988; Shields, 1962; Tellegen et al., 1988), and from studies of twin pairs reared together (Eaves, Eysenck, & Martin, 1989; Rose & Kaprio, 1988; Rose, Koskenvuo, Kaprio, Sarna, & Langinvainio, 1988) and of twin pairs and their parents, siblings, and adult children (Eaves, Heath, Neale, Hewitt, & Martin, 1992) all support the conclusion that family resemblance for such personality traits as Extraversion and Neuroticism, assessed in adolescence or adulthood, is strongly influenced by genetic factors and that there is no significant influence of shared family environment. This evidence for a genetic contribution to personality differences, however, relies almost entirely on studies using self-report assessments of personality. A multimethod approach to personality assessment, combining self-report data with ratings by informants and objective test measures (e.g., Campbell & Fiske, 1959; Cattell, 1957; Eysenck, 1960), has certain advantages: Objective performance measures are less vulnerable to attempts to "fake" (e.g., Furnham, 1986) responses; rating measures may prove less vulnerable to the "state" effects observed when self-report instruments are used with clinical populations (Coppen & Metcalfe, 1965; Dodwell, 1988; Hirschfeld et al., 1983; Knowles & Kreitman, 1965). Evidence for a genetic influence on personality would be more convincing if confirmed using multiple modes of assessment.

Large sample sizes are required to ensure adequate statistical power for resolving genetic and environmental contributions to personality differences or other continuously distributed traits (Heath, Kendler, Eaves, & Markell, 1985; Martin, Eaves, Kearsey, & Davies, 1978). This necessarily limits the feasibility of using objective test data but need not be a problem for ratings by informants (*ratings*). Obtaining both self-report and reciprocal rating data on family members (e.g., Bem & Allen, 1974; Condon, 1988; Crandall, 1976; Eaves & Last, 1980; Heath, Berg, et al., 1985; Neale & Stevenson, 1989; Silberg et al., 1991) provides a useful check on the validity of the self-reports and permits a resolution of the contributions of the personality of the rater, as well as of the personality of the individual being rated (the *ratee*), to ratings. In the present study, we analyzed data on adult female twin pairs who provided both self-ratings and ratings of their co-twin, and we examined the consistency of evidence for genetic effects on personality when self-report and reciprocal rating data were combined.

Method

Sample

Data for this article came from an ongoing study of risk factors for common psychiatric disorders using female same-sex twins from the

Data collection was supported by Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) Grant MH-40828. Data analysis was also supported by ADAMHA Grants AA03539, AA07535, AA07728, and DA05588. The Virginia Twin Registry was established and is maintained by W. E. Nance and L. Corey, and it is supported by National Institutes of Health Grants HD-26746 and NS-25630.

We are grateful to Michael Hodge for assistance with data analyses. We are grateful to Kay Phillips and to an anonymous reviewer for insightful comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to A. C. Heath, Department of Psychiatry, Washington University School of Medicine, 4940 Audubon Avenue, St. Louis, Missouri 63110.

Virginia Twin Register, a population-based register of all multiple births in Virginia from 1915 onward. Current addresses were obtained by matching to state records and by similar tracing methods. In 1987 and 1988, approximately 3,300 questionnaires were mailed to members of White female same-sex pairs born between 1935 and 1971, at their last known address. Usable data for the present article were returned by 2,114 twins (64% individual response rate), including 826 complete pairs. Zygosity of a twin pair was assigned on the basis of responses to questions about physical similarity and about having been confused with one another during childhood, a method which has been found to agree with the results of blood-typing in 95% of cases (Eaves et al., 1989; Kasriel & Eaves, 1976). In some cases these data were supplemented by information about blood groups and a review of current photographs (Kendler, Kessler, Heath, Neale & Eaves, in press). This algorithm identified 460 monozygotic pairs and 366 dizygotic pairs for whom we have personality data.

Measure

A self-report rating instrument was developed using items from the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975) that had high loadings on Extraversion and Neuroticism in a factor analysis of Australian data (Heath, Jardine, Eaves, & Martin, 1988) and the wording of which converted easily to a rating format. A 12-item Neuroticism scale and an 8-item Extraversion scale were created (see Results section for wording of the individual items). The regular instructions and layout for the EPQ were modified, with parallel columns provided for self-reports and for ratings of the co-twin. Respondents were instructed not to consult with their twin; they were asked first to answer questions about themselves and then to answer the same questions about their co-twin.

Self-report and co-twin rating scores were derived as the proportion of responses checked *yes* separately for Extraversion and Neuroticism scales. Nonnormality of the joint distribution of twin pairs was tested by computing linear, quadratic, and cubic terms for the polynomial regression of intrapair variance on twin pair means (Heath, Neale, Hewitt, Eaves, & Fulker, 1989; Jinks & Fulker, 1970) for each scale. All regressions were modest and were not reduced by standard (e.g., Eaves et al., 1989) logarithmic, square root, or arcsine transformations of the data, so that no further data transformation was required. Internal consistency of the two scales was high both for self-report (Cronbach's αs = .83 for Extraversion [N = 2,108] and .83 for Neuroticism [N = 2,107]) and for co-twin ratings (.83 for Extraversion [N = 2,098] and .80 for Neuroticism [N = 2,096]). A subset of subjects in this study had previously completed a conventional self-report short-form EPQ-Revised (EPQ-R), which included 12-item Extraversion and 12-item Neuroticism scales (Eysenck, Eysenck, & Barrett, 1985; see Eaves et al., 1992). Correlations between self-report scores on the modified instrument and the conventional instrument were .78 for Extraversion (N = 1,240) and .73 for Neuroticism (N = 1,241). Correlations between ratings of the respondent by her co-twin on the modified instrument and the twin's own self-report on the EPQ-R were .56 for Extraversion (N = 1,132) and .46 for Neuroticism (N = 1,135). There thus appears to be reasonable consistency of responses despite the change in questionnaire format.

Sampling Representativeness

Nonrepresentative sampling of twin pairs with respect to personality traits could seriously bias estimates of genetic and environmental parameters (Lykken, McGue, & Tellegen, 1987; Martin & Wilson, 1982; Neale, Eaves, Hewitt, & Kendler, 1989). However, most plausible forms of nonrandom sampling can lead to differences in mean and variance of personality scores between monozygotic and dizygotic pairs, differences that are not predicted in most genetic models (Heath et al., 1989). In testing for differences in mean and variance, we have ignored the nonindependence of observations on twin pairs. Our significance tests will thus be conservative; namely, they will tend to overestimate the significance of mean or variance differences (McGue, Wette, & Rao, 1989). Monozygotic twins gave higher Neuroticism ratings of their co-twins (M = 0.42) than did dizygotic twins (M = 0.40), but this difference, although statistically significant, $t(2{,}081) = 2.24$, $p < .05$, was slight. No other statistically significant zygosity differences in mean or variance were found, but there was a trend for the variance of Extraversion self-reports and co-twin ratings to be reduced in monozygotic twin pairs, $F(985, 1{,}106) = 1.13$, $p = .05$; $F(979, 1{,}100) = 1.10$, $p = .14$. This decreased monozygotic variance would also be predicted under reciprocal inhibitory environmental interaction between twins, occurring against a background of genetic effects on personality (Carey, 1986; Eaves, 1976; see below). With most forms of nonrandom sampling, one would expect to find a decreased, rather than an increased, variance for dizygotic twins.

Data from twins whose co-twin refused to cooperate in the study provided an additional check on response biases (Heath et al., 1989): In the absence of a sampling bias, the means and variances of Extraversion and Neuroticism scores for this group are not expected to differ significantly from those of twins from pairs in which both members returned questionnaires. Similarly, the means and variances of the Extraversion and Neuroticism ratings of those twins who did not themselves participate in the study (provided by their cooperative co-twins) are not expected to differ from the ratings of twins who did participate. Dizygotic twins who did not participate in the study received significantly elevated Extraversion ratings from their co-twins (M = 0.68, compared with M = 0.63 for dizygotic twins from concordant cooperative pairs), $t(978) = 2.24$, $p < .05$, but no significant difference was found for self-reports by the cooperative dizygotic twins from concordant pairs (M = 0.64) compared with the dizygotic twins whose co-twin refused to cooperate (M = 0.62), $t(984) = 0.87$, $p > .05$. No significant differences were found for Extraversion in monozygotic twins or for Neuroticism in twins from either zygosity group. No significant differences in variance were found for any group. Any sampling bias, and consequent bias to estimates of genetic and environmental parameters, must therefore be very slight.

Data Analysis

Factor analysis was performed to confirm the a priori assignment of items to Extraversion and Neuroticism scales, separately for the self-report data and the co-twin rating data. Tetrachoric correlations between item responses were computed for the entire sample using PRELIS (Jöreskog & Sörbom, 1986). A two-factor model was fitted by the method of maximum likelihood, and factor loadings were rotated to an oblique solution using promax criteria. Observations on members of a twin pair are not independent. However, provided that sampling was random with respect to Extraversion and Neuroticism traits (and, from our checks on sampling representativeness, this appears to be at least approximately the case), this nonindependence of observations should not bias estimates of factor loadings (although it would lead to underestimation of the standard errors of those loadings).

The structural equation model used for model fitting is summarized in the form of a path diagram in Figure 1. The model has two components: the genetic model, specifying the genetic and environmental determinants of the "true" phenotype (e.g., Heath et al., 1989), and the reciprocal rating model, specifying the relationships between the true phenotypes of the twin pairs and their self-reports and co-twin ratings (e.g. Heath, Berg, et al., 1985; Neale & Stevenson, 1989; Silberg et al., 1991). The full genetic model allowed for the contributions to personality differences of additive gene action (path coefficient h in Figure 1), genetic dominance or epistasis (d: these two forms of genetic nonadditivity are confounded in twin data, unless data on other relatives are also available), shared family environment (c: representing those environmental influences that are shared by twins reared in the same fam-

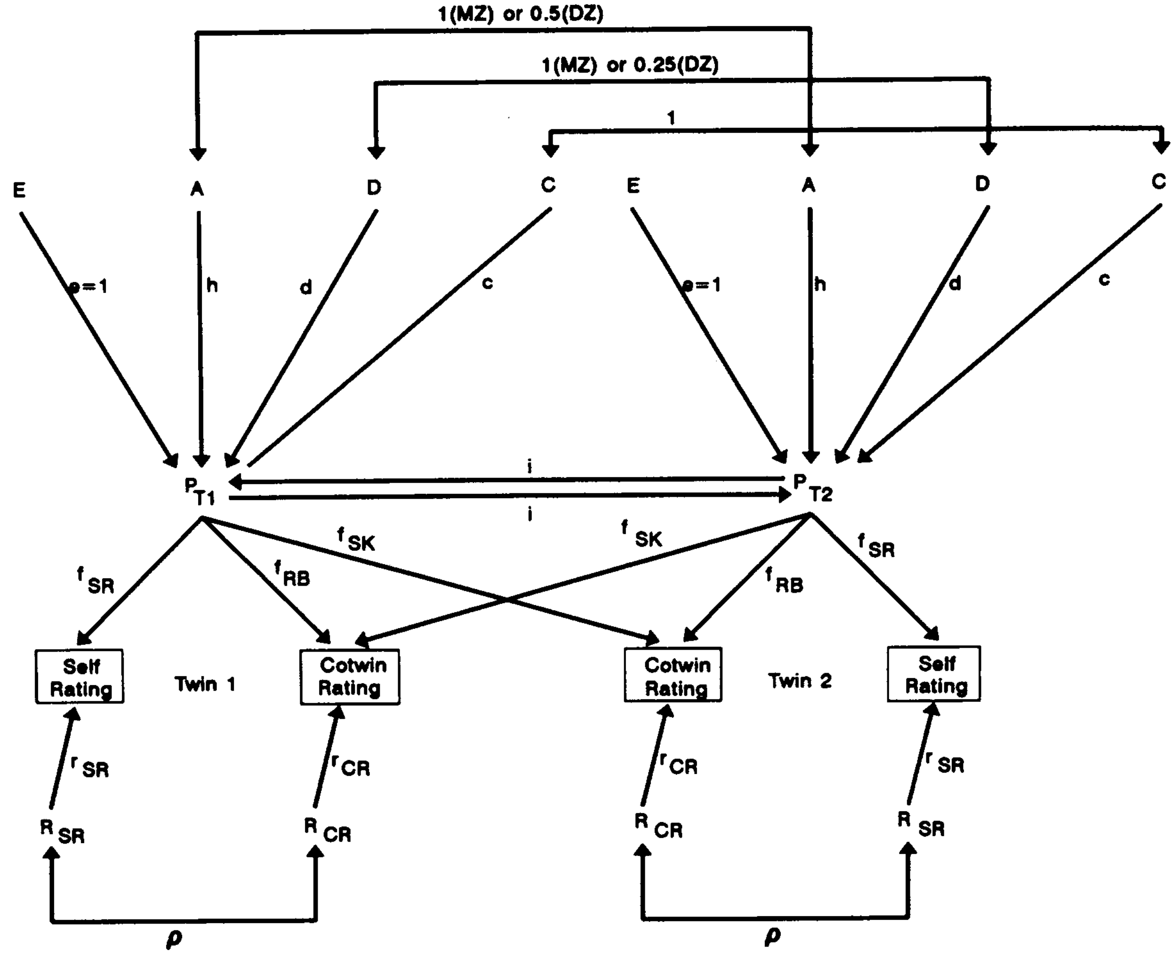

Figure 1. Reciprocal rating model for twin data. (MZ = monozygotic; DZ = dizygotic; P = true phenotype; E = nonshared environmental deviation; A = additive genetic deviation; D = nonadditive genetic [dominance or epistatic] deviation; C = shared environmental deviation, with path coefficients e, h, d, and c representing the path regressions of P on E, A, D, and C, respectively. T1 = 1st twin; T2 = 2nd twin; i = sibling interaction effect. R_{SR} and R_{CR} are residual [measurement error] effects on self-rating and co-twin rating [with corresponding path coefficients r_{SR} and r_{CR} and correlation ρ]. Paths from true phenotype to self-rating, rating of co-twin [rater bias parameter], and rating by co-twin [sibling knowledge parameter] are denoted by f_{SR}, f_{RB}, and f_{SK}, respectively.)

ily), nonshared environment (*e*), and the reciprocal environmental influence of the true personality of each twin on that of the co-twin (*i*: sometimes called *sibling interaction effects*; Carey, 1986; Eaves, 1976; Heath et al., 1989). Genetic and environmental paths were constrained to be equal in first and second twins and in monozygotic and dizygotic twin pairs. However, the correlation between the underlying additive genetic and nonadditive genetic factors of first and second twins from a pair, from genetic theory (see, e.g., Bulmer, 1980), will be unity for monozygotic twin pairs and .5 and .25, respectively, for dizygotic twin pairs; that between shared environmental deviations will be unity in each case. (The correlation between the additive genetic deviations of dizygotic twins is predicted to be .5 only if there is random mating for personality, but data on large samples of spouse pairs confirm a zero marital correlation for both Extraversion and Neuroticism: Eaves et al., 1989; Eaves et al., 1992). In data on twin pairs reared together, the effects of genetic nonadditivity and shared environment are confounded (Martin et al., 1978), with genetic nonadditivity tending to reduce the dizygotic correlation to less than one half the corresponding monozygotic correlation and shared environment tending to increase the dizygotic correlation to greater than one half the monozygotic correlation. Because analyses of other data sets in which these parameters are not confounded have given no evidence of shared environmental effects on Extraversion and Neuroticism (see the introduction), we have set the shared environmental parameter (*c*) to zero, except in the case in which we have specifically tested a nongenetic model. The effects of genetic nonadditivity and of inhibitory reciprocal environmental interaction (i.e., if Extraversion in one twin tends to inhibit Extraversion in the co-twin, and vice versa), although not completely confounded in twin data, are nonetheless extremely difficult to distinguish from each other. Reciprocal inhibitory interaction can generate very low dizygotic twin correlations, or even negative twin correlations, that cannot be explained by nonadditive genetic effects (Carey, 1986; Eaves, 1976), but over a range of parameter values, the power of twin data to distinguish genetic nonadditivity and sibling environmental interaction is very low (e.g., Jardine, 1985).

The reciprocal rating component of the model assumes that self-report personality is determined by the individual's true underlying personality (the *self-report* parameter f_{SR}) plus measurement error (r_{SR}).

Co-twin ratings (i.e., ratings by the respondent of the co-twin's personality), however, are allowed to be influenced by the true personality of the ratee (the *sibling knowledge* parameter f_{SK}) and by the personality of the rater (the *rater bias* parameter f_{RB}) as well as by measurement error of the co-twin rating (r_{CR}). The sibling knowledge parameter is thus a measure of the extent to which a respondent is successful in assessing the personality of her co-twin. The rater bias parameter represents a generalized tendency of a rater to give ratings of the behaviors of others that are influenced by her own personality, for example, a tendency for more extraverted raters to underestimate the Extraversion of those they are rating. The model also allows for correlations between measurement errors for self-report and for ratings of the co-twin (ρ), such as might arise if some respondents have a tendency to endorse *yes* items both for self-reports and for co-twin ratings (a tendency that might be exacerbated by the questionnaire format of parallel columns for self-report and for ratings of the respondent's co-twin). In some analyses, we have allowed either the sibling knowledge parameter (f_{SK}) or the measurement error parameters for ratings of the co-twin (r_{CR} or ρ) to depend on zygosity. This allows for the possibility that monozygotic twins are better able to assess the personalities of their co-twins than are dizygotic twins. It is important to note that our rater bias model is modeling variances and covariances, not means. We show below that for some items there are also mean differences between self-report and co-twin ratings that are almost certainly a rater bias effect, but we do not attempt to model these mean effects.

When the genetic and reciprocal rating components of the model are combined, it is not possible to estimate the absolute magnitudes of paths h, e, and d (or c), only their relative magnitudes. We therefore fixed the value of the parameter e to unity and estimated h and d (or c) as free parameters.

Covariance matrices were computed separately for each zygosity group, giving the variances and covariances of the self-report and co-twin ratings of first and second twins. Twins were assigned as first or second members of a pair on the basis of birth order, when this information was available, or at random otherwise. Because Extraversion and Neuroticism scores were almost uncorrelated, separate analyses were conducted for each trait. Models were fitted to the twin covariance matrices by the method of maximum likelihood using LISREL (Jöreskog & Sörbom, 1983), in a multiple-group analysis (Heath et al., 1989; Neale, Heath, Hewitt, Eaves, & Fulker, 1989). This provided an overall chi-square test of goodness of fit of each model. The goodness of fit of nested models was compared by a likelihood ratio chi-square test. Parameter estimates reported for best fitting models have been standardized to unit variance. In the case of models involving reciprocal sibling environmental interaction, which therefore predict differences in variance between monozygotic and dizygotic twins, we have standardized parameters of the genetic model to unit variance in monozygotic pairs. For such models, we report the proportion of the total phenotypic variance (VP) attributable to both the direct genetic effect on the phenotype and the indirect effect that is mediated through the environmental impact of the co-twin's phenotype (*genotype–environment correlation*). Our estimate of the total genetic variance is

$$\frac{h^2}{VP(1-i)^2}.$$

For comparisons of parameters of the reciprocal rating component of the model, we also restandardized dizygotic parameters to unit variance.

Results

Endorsement frequencies for the modified EPQ items are presented in Table 1. For most items, endorsement frequencies were very similar for self-report and for co-twin ratings. However, respondents were more likely to rate themselves than to rate their co-twins as rather nervous and moody ("sometimes feels 'just miserable' "), sensitive ("feelings easily hurt"), and worriers ("worries too long after an embarrassing experience" and "often troubled about feelings of guilt"). Respondents were more likely to rate their co-twins than to rate themselves as able to enliven parties and as irritable.

Also shown in Table 1 are the factor loadings obtained in separate factor analyses of the self-report and co-twin rating data. In each analysis, factor loadings after rotation clearly identified separate Neuroticism and Extraversion factors. Interfactor correlations in an oblique solution were modest ($r = -.25$ for self-report; $r = -.24$ for co-twin ratings). Factor loadings for the Extraversion factor were comparable for self-report and for co-twin rating data; those for the Neuroticism factor were slightly higher for the self-report data. Two items relating to the ability to enliven parties, which were almost tautologous (Items 12 and 20), had very high loadings on the Extraversion factor both in the self-report and in the co-twin rating data.

Table 2 presents variance–covariance matrices for the self-reports and co-twin ratings of first and second twins, separately for Extraversion and Neuroticism, and for monozygotic and dizygotic twin pairs. Agreement between a twin's self-report and how she was rated by her co-twin was somewhat higher for Extraversion (correlations ranging from .54 to .63) than for Neuroticism (.45 to .51) and was somewhat better for monozygotic pairs (.49 to .63) than for dizygotic pairs (.45 to .59). Monozygotic twin correlations (r_{MZ}) for Extraversion and for Neuroticism were higher than the corresponding dizygotic correlations (r_{DZ}), whether these were based on self-report data (Extraversion: $r_{MZ} = .39$, $r_{DZ} = -.04$; Neuroticism: $r_{MZ} = .35$, $r_{DZ} = .10$) or on ratings by the respondent's co-twin (Extraversion: $r_{MZ} = .27$, $r_{DZ} = -.17$; Neuroticism: $r_{MZ} = .30$, $r_{DZ} = .14$).

Neuroticism Ratings

The results of fitting the reciprocal rating model of Figure 1 to the twin covariance matrices for Neuroticism are summarized in Table 3. When all parameters of the reciprocal rating model were constrained to be equal across zygosity groups, the full model, setting $c = 0$ (i.e., assuming no shared environmental effects: Model 1), gave a very poor fit by chi-square goodness-of-fit test. Submodels that dropped the genetic dominance parameter (Model 3) or dropped a reciprocal sibling interaction parameter, which allowed for the inhibitory effects of one twin's Neuroticism on that of the co-twin (Model 2), did not give a significantly worse fit than that of the full model, by a likelihood ratio chi-square test ($\chi^2_1 = 2.32$, $p = .13$; $\chi^2_1 = 0.14$, $p = .71$). However, a model that dropped both dominance and sibling interaction effects (Model 5) did give a significantly worse fit ($\chi^2_2 = 6.61$, $p < .05$), suggesting that at least one of these parameters was needed in the model. Because our estimate of the sibling interaction parameter was close to zero in the full model, we retained the genetic dominance parameter (Model 2). A nongenetic model was also fitted (Model 4), in which additive genetic and dominance effects were fixed to zero and a shared environmental parameter (c) was included, but this gave a substantially worse fit than the full model ($\chi^2_2 = 47.40$, $p < .001$).

Table 1
Endorsement Frequencies of Extraversion and Neuroticism Items and Maximum Likelihood Factor Loadings in Two-Factor Oblique Solution

	Endorsement (%)		Factor loading			
			Self-report		Rating co-twin	
Item	Self-report	Rating co-twin	N	E	N	E
1. Mood often goes up and down	51.6	52.1	.70	.07	.67	.11
2. Rather talkative	61.2	61.1	.11	.62	.11	.60
3. Sometimes feels "just miserable" for no reason	48.0	41.1	.63	−.01	.62	.02
4. Rather lively	73.2	69.1	−.09	.67	−.09	.66
5. Irritable	36.1	43.9	.68	.05	.62	.11
6. Enjoys meeting new people	79.6	77.3	−.18	.61	−.16	.58
7. Feelings easily hurt	68.9	62.0	.51	−.10	.45	−.15
8. Can usually enjoy a lively party	75.3	75.5	−.09	.71	−.09	.74
9. Often feels fed up	44.5	46.3	.67	−.03	.61	.01
10. Usually takes initiative in making new friends	60.5	60.3	−.05	.67	−.05	.67
11. Rather nervous	39.5	31.7	.76	−.04	.73	−.08
12. Can easily put some life into a dull party	43.2	50.3	.14	.97	.15	.97
13. A worrier	66.7	53.1	.67	−.07	.61	−.20
14. Tense or highly strung	37.0	34.1	.77	.06	.75	.08
15. Likes mixing with people	71.6	72.0	−.15	.70	−.17	.67
16. Worries too long after an embarrassing experience	56.8	38.8	.44	−.19	.43	−.28
17. Suffers from nerves	28.5	25.2	.80	.06	.75	.03
18. Often feels lonely	37.5	33.7	.64	−.06	.60	−.06
19. Often troubled about feelings of guilt	40.0	25.7	.58	−.02	.52	−.09
20. Can get a party going	40.5	49.2	.09	.97	.13	.96

Note. N = Neuroticism; E = Extraversion.

Thus, Model 2 was the simplest model consistent with the data. We did not attempt to fit a model allowing for additive genetic and shared environmental effects (*h* and *c*) because, in the presence of genetic dominance, this would yield a zero estimate for the parameter *c*, resulting from the confounding of dominance and shared environmental effects in twin data (Martin et al., 1978; Heath et al., 1989).

An attempt to simplify the reciprocal rating component of the model, dropping the rater bias parameter f_{RB} from Model 2 (Model 6), did not lead to a deterioration in fit ($\chi^2_1 = 0.34$, $p = .56$), implying that there is no tendency for individuals with high Neuroticism levels to be more (or less) likely to rate others as neurotic. Fixing the path f_{SK} to zero (Model 7), which implies that twins have no knowledge of the personalities of their co-twins, did lead to a significant worsening of fit, as did imposition of the constraint that residual effects on self-report and co-twin ratings be uncorrelated (Model 8). Thus, of those models that constrained the parameters of the reciprocal rating model to be equal across zygosity groups, Model 6 was the best fitting model.

To explore the reasons for the poor fit of Model 6, we tried relaxing the constraint that reciprocal rating parameters be equal across zygosity groups. A model that allowed for zygosity differences in residual variance for co-twin ratings, and in the correlation between residual effects on self-report and on ratings of the co-twin, gave an excellent fit to the data (Model 9: $\chi^2_{11} = 7.40$, $p = .77$) and resulted in a substantial improvement in fit compared with Model 6 ($\chi^2_3 = 38.02$, $p < .001$). Allowing for zygosity differences in knowledge of the co-twin's personality (Model 10) did not produce any improvement in fit over Model 6 ($\chi^2_1 = 1.42$, $p = .23$), and a model which allowed for both zygosity-dependent sibling knowledge and residuals (Model 11) did not improve on the fit of Model 9 ($\chi^2_1 = 0.19$, $p = .66$). Deleting the genetic dominance parameter from Model 9 (Model 12) did not lead to a significant deterioration in fit ($\chi^2_1 = 0.38$ $p = .54$), and this reduced model gave an excellent fit to the data. Equating the residual variances for self-report and for co-twin rating data ($r_{SR} = r_{SR(DZ)} = r_{CR} = r_{CR(DZ)} = r_{CR(DZ)}$), although still allowing for zygosity differences in the correlation between residual effects (Model 13), produced a still simpler model, whose fit was no worse than that of Model 12 ($\chi^2_2 = 0.11$, $p = .95$). A nongenetic model with zygosity-dependent residual correlation still gave a very poor fit to the data (Model 14). Thus, Model 13 was the simplest model consistent with the observed Neuroticism data. Standardized parameter estimates for the Neuroticism data in this model are summarized in the form of a path diagram in Figure 2. (We note that although parameters r_{SR} and r_{CR} are assigned the same unstandardized value in Model 13, their standardized values are not expected to be the same.)

Table 2
Covariance Matrices for Self-Reports and Ratings of Co-Twin, for Extraversion and Neuroticism

Twin and measure	Neuroticism 1	2	3	4	Extraversion 1	2	3	4
Monozygotic (MZ) female pairs								
Twin 1								
1. Self-report	0.792	**0.599**	**0.350**	**0.491**	0.862	**0.414**	**0.390**	**0.616**
2. Rating of co-twin	0.464	0.758	**0.505**	**0.299**	0.364	0.893	**0.628**	**0.273**
Twin 2								
3. Self-report	0.280	0.396	0.811	**0.531**	0.354	0.581	0.956	**0.371**
4. Rating of co-twin	0.371	0.221	0.406	0.719	0.578	0.261	0.366	1.021
Dizygotic (DZ) female pairs								
Twin 1								
1. Self-report	0.768	**0.156**	**0.098**	**0.449**	0.997	**−0.007**	**−0.040**	**0.542**
2. Rating of co-twin	0.116	0.718	**0.461**	**0.142**	−0.008	1.073	**0.586**	**−0.171**
Twin 2								
3. Self-report	0.077	0.349	0.799	**0.213**	−0.040	0.608	1.002	**−0.130**
4. Rating of co-twin	0.320	0.098	0.155	0.661	0.565	−0.186	−0.136	1.091

Note. Correlations are in boldface in the upper triangle of each matrix. For MZ, *ns* = 453 and 450 pairs for Neuroticism and Extraversion, respectively. For DZ, *ns* = 362 and 361 pairs for Neuroticism and Extraversion, respectively.

Extraversion Ratings

Table 4 summarizes the results of fitting models to the Extraversion reciprocal rating data. Once again, considering first those models that constrain parameters of the reciprocal rating model to be equal across zygosity groups, the full model (with $c = 0$: Model 1) was rejected by chi-square test of goodness of fit. Deleting a genetic dominance parameter (Model 3) did not lead to a significant worsening of the fit compared with that of the full model, by a likelihood ratio chi-square test ($\chi^2_1 = 3.21$, $p = .07$), whereas deleting a sibling interaction parameter, which

Table 3
Results of Model Fitting: Neuroticism

Constraints and model no.	Goodness of fit *df*	χ^2	*p*
Without zygosity-dependent parameters			
1. Full model	11	44.94	<.001
2. No sibling interaction ($i = 0$)	12	45.08	<.001
3. No dominance ($d = 0$)	12	47.26	<.001
4. No genetic effects ($h = d = i = 0; c > 0$)	13	92.34	<.001
5. No dominance, no sibling interaction ($d = i = 0$)	13	51.55	<.001
6. No sibling interaction, no rater bias ($i = f_{RB} = 0$)	13	45.42	<.001
7. No sibling interaction, no sibling knowledge ($i = f_{SK} = 0$)	13	275.16	<.001
8. No sibling interaction, no rater bias, uncorrelated residuals ($i = \rho = f_{RB} = 0$)	14	131.10	<.001
With zygosity-dependent parameters			
9. Residual ($r_{CR} \neq r_{CR(DZ)}, \rho_{MZ} \neq \rho_{DZ}, i = f_{RB} = 0$)	11	7.40	.77
10. Sibling knowledge ($f_{SK} \neq f_{SK(DZ)}; i = f_{RB} = 0$)	12	44.00	.001
11. Residual, sibling knowledge ($r_{CR} \neq r_{CR(DZ)}; \rho_{MZ} \neq \rho_{DZ}; f_{SK} \neq f_{SK(DZ)}, i = f_{RB} = 0$)	10	7.21	.71
12. Residual, no dominance ($r_{CR} \neq r_{CR(DZ)}; \rho_{MZ} \neq \rho_{DZ}; d = i = f_{RB} = 0$)	12	7.78	.80
13. Correlation between residuals ($\rho_{MZ} \neq \rho_{DZ}; r_{CR} = r_{CR(DZ)} = r_{SR} = r_{SR(DZ)}; d = i = f_{RB} = 0$)	14	7.89	.90
14. Residual, no genetic effects ($\rho_{MZ} \neq \rho_{DZ}; h = d = i = f_{RB} = 0; c > 0$)	14	23.62	.05

Note. i = sibling interaction; d = dominance; h = heritability; c = shared environment; RB = rater bias; SK = sibling knowledge; CR = co-twin rating; DZ = dizygotic; MZ = monozygotic.

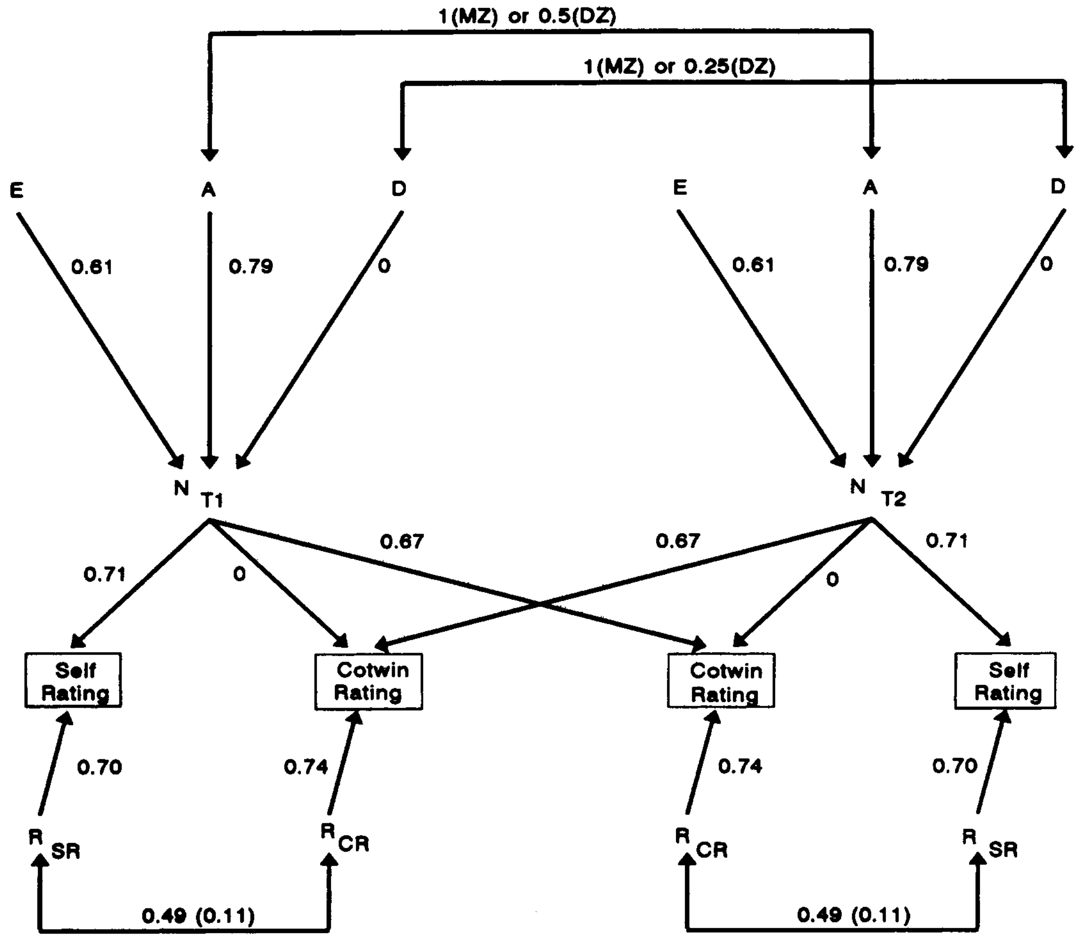

Figure 2. Standardized parameter estimates for reciprocal rating model: Neuroticism. (Parameters specific to dizygotic pairs are given in parentheses in figure. MZ = monozygotic; DZ = dizygotic; E = nonshared environmental deviation; A = additive genetic deviation; D = nonadditive genetic [dominance or epistatic] deviation; N = Neuroticism; T1 = 1st twin; T2 = 2nd twin. R_{SR} and R_{CR} are residual [measurement error] effects on self-rating and co-twin rating.)

allowed for the reciprocal inhibitory effects of each twin's Extraversion on that of the co-twin (Model 2), did lead to a significantly worse fit ($\chi_1^2 = 6.87$, $p < .01$). When we tested a nongenetic model, which included a shared environmental path (*c*) but set additive genetic and dominance paths to zero (Model 4), the model gave a very poor fit to the data. Comparison of the first four models thus identified Model 3 as the simplest model consistent with the data.

Attempts to simplify the reciprocal rating component of the model by dropping from Model 3 the path f_{SK} from twin's true phenotype to co-twin's rating (i.e., assuming that twins have no knowledge of each other's personalities: Model 6), the path f_{RB} from twin's true phenotype to twin's rating of co-twin (i.e., assuming no rater bias: Model 5), or the correlation between residual effects (Model 7) in all cases led to significantly poorer fits than that obtained with Model 3 ($\chi_1^2 = 374.52$, $p < .001$; $\chi_1^2 = 6.32$, $p < .05$; $\chi_1^2 = 11.40$, $p < .001$, respectively). In contrast to our findings for Neuroticism, it thus appears that ratings of Extraversion are influenced by the personality of the rater (rater bias), as well as the personality of the ratee.

As in the case of Neuroticism, we explored the reasons for the poor fit of Model 3 by relaxing the constraint that rating parameters be equal across zygosity groups. Compared with the goodness of fit of Model 3, allowing for zygosity differences in the residual variance of the co-twin ratings and in the correlation between residual effects on the self-report and co-twin rating (Model 8: $\chi_2^2 = 8.68$, $p = .01$), or in the sibling knowledge parameter (Model 9: $\chi_1^2 = 3.99$, $p < .05$), in each case led to a significant improvement in fit. A model that allowed both for zygosity differences in the residual variance of co-twin ratings and the correlation between residuals, and for zygosity differences in sibling knowledge (Model 10), gave a significantly better fit than Models 8 and 9 (likelihood ratio chi-squares: $\chi_1^2 = 10.35$, $p = .001$; and $\chi_2^2 = 15.04$, $p < .001$, respectively) and gave a good fit to the data ($p = .17$). Constraining the self-report and co-twin residual variances to be equal, and to be equal across zygosity groups, while estimating different correlations between the residuals, led to a significant worsening of fit (Model 11: $\chi_2^2 = 12.28$, $p = .001$). Constraining the correlation between residuals to be equal across zygosity groups, while estimating separate

Table 4
Results of Model Fitting: Extraversion

Constraints and model no.	df	Goodness of fit χ^2	$p \leq$
Without zygosity-dependent parameters			
1. Full model	11	28.58	.003
2. No sibling interaction (i = 0)	12	35.45	.001
3. No dominance (d = 0)	12	31.79	.001
4. No genetic effects ($h = d = i = 0$; $c > 0$)	12	102.81	.001
5. No dominance, no rater bias ($d = f_{RB} = 0$)	13	38.11	.001
6. No dominance, no sibling knowledge ($d = f_{SK} = 0$)	13	406.31	.001
7. No dominance, uncorrelated residuals ($d = \rho = 0$)	13	43.19	.001
With zygosity-dependent parameters			
8. Residual, residual correlation ($r_{CR} \neq r_{CR(DZ)}$; $\rho_{MZ} \neq \rho_{DZ}$; $d = 0$)	10	23.11	.01
9. Sibling knowledge ($f_{SK} \neq f_{SK(DZ)}$; $d = 0$)	11	27.80	.003
10. Sibling knowledge, residual, residual correlation[a] ($f_{SK} \neq f_{SK(DZ)}$; $r_{CR} \neq r_{CR(DZ)}$; $\rho_{MZ} \neq \rho_{DZ}$; $d = 0$)	9	12.76	.17
11. Sibling knowledge, residual correlation ($f_{SK} \neq f_{SK(DZ)}$; $\rho_{MZ} \neq \rho_{DZ}$; $r_{CR} = r_{CR(DZ)}$; $d = 0$)	11	25.45	.008
12. Sibling knowledge, residual ($f_{SK} \neq f_{SK(DZ)}$; $r_{CR} \neq r_{CR(DZ)}$; $\rho_{MZ} = \rho_{DZ}$; $d = 0$)	10	13.04	.22
13. Sibling knowledge, residual, residual correlation[a], no genetic effect ($f_{SK} \neq f_{SK(DZ)}$; $r_{CR} \neq r_{CR(DZ)}$; $\rho_{MZ} \neq \rho_{DZ}$; $h = d = i = 0$; $c > 0$)	10	59.45	.001

Note. i = sibling interaction; d = dominance; h = heritability; c = shared environment; RB = rater bias; SK = sibling knowledge; CR = co-twin rating; SR = self-rating; DZ = dizygotic; MZ = monozygotic.
[a] Correlation between residuals R_{SR} and R_{CR}.

residual variances for co-twin ratings by monozygotic versus dizygotic twins (Model 12), gave an excellent fit to the data, one that was no worse than that of Model 10 ($\chi^2_2 = 0.28$, $p = .95$). A nongenetic model with zygosity-dependent sibling knowledge and residual parameters (Model 13) still gave a very poor fit to the data. Model 12 was therefore the most parsimonious model that was able to account for the Extraversion data. Figure 3 summarizes in the form of a path diagram standardized parameter estimates in Model 12 for monozygotic twin pairs. (Under reciprocal sibling interaction [e.g., Eaves, 1976; Carey, 1986], the variances of the true Extraversion scores, and therefore standardized parameter estimates, are predicted to differ between monozygotic and dizygotic twin pairs.) It should be noted that, because of the negative genotype–environment correlation that arises under sibling interaction when $i < 0$, $h^2 + e^2$ is not expected to be equal to unity. Estimates of the parameters of the reciprocal rating model for dizygotic twin pairs, restandardizing to unit variance, were $f_{SR} = .82$; $f_{RB} = -.11$; $f_{SK} = .69$; $r_{SR} = .57$; $r_{CR} = .71$; and $\rho_{DZ} = .14$.

Discussion

The evidence for a genetic influence on personality derives largely from studies using self-report measures. In those cases in which objective test data, or data that were based on ratings by an informant, have been used, sample sizes have been too low to permit a powerful resolution of genetic and nongenetic hypotheses (Martin et al., 1978). To test whether a genetic influence can also be observed in data that are based on informant ratings, we analyzed reciprocal ratings of Extraversion and Neuroticism in a sample of 826 adult female twin pairs. By combining self-report and co-twin rating data in a single analysis, we have attempted to estimate and statistically control for the impact of biases in self-perception and in perception of others.

Accuracy of Personality Ratings

Our reciprocal rating instrument, modified from the short-form EPQ-R (Eysenck et al., 1985), had acceptable internal consistencies, which were comparable for self-report and for co-twin rating formats, for both Extraversion and Neuroticism (coefficient αs = .8–.85). We found good agreement between self-report personality measures and ratings by the respondent's co-twin. Endorsement frequencies were similar for self-report and co-twin rating formats for most items, although respondents were more likely to rate themselves as worried and their co-twins as irritable! (In a study of high-school age twin pairs, in which respondents were asked to make comparative ratings of themselves against their co-twins, Loehlin and Nichols [1976] also found that twins tended to attribute more self-confidence to their co-twin.) Factor structures were also similar for self-report and for co-twin rating data. Observed correlations between self-report Extraversion and Neuroticism and ratings provided by the co-twin were moderately high (.45–.63). However, because self-report data are an imperfect reflection of an individual's true underlying personality, these raw correlations will underestimate the accuracy of the personality ratings.

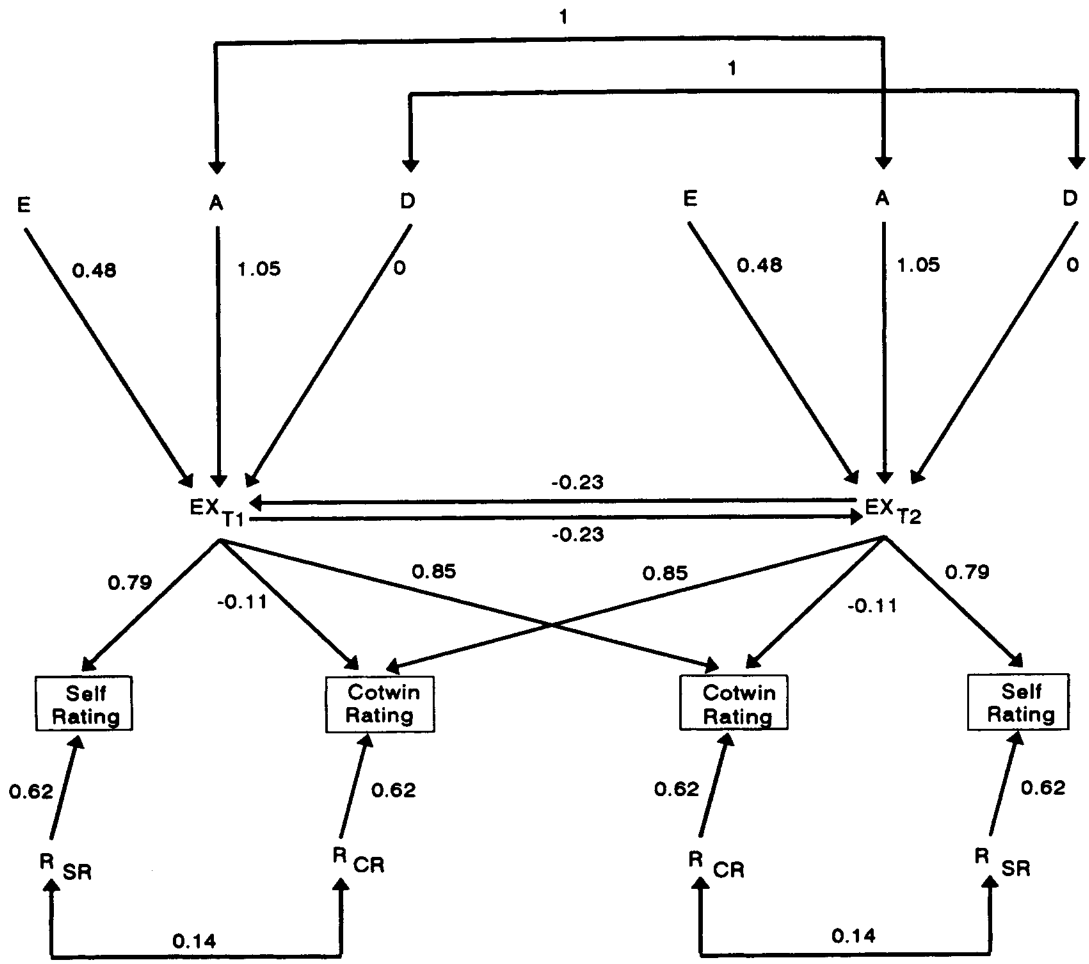

Figure 3. Standardized parameter estimates for monozygotic twin pairs in best fitting reciprocal rating model: Extraversion. (E = nonshared environmental deviation; A = additive genetic deviation; D = nonadditive genetic [dominance or epistatic] deviation. EX = Extraversion; T1 = 1st twin; T2 = 2nd twin. R_{SR} and R_{CR} are residual [measurement error] effects on self-rating and co-twin rating.)

Reciprocal rating data were analyzed under a model that assumed that self-report personality scores were a function of the underlying true personality score plus measurement error, but that allowed personality ratings by the respondent of her co-twin to be a function of the true personality of the respondent (rater bias), the true personality of the person being rated (sibling knowledge), and measurement error (Figure 1). Additionally, the model allowed measurement errors to be correlated between self-report and co-twin rating scores. Because self-report and co-twin ratings were made on the same occasion and because we used parallel columns for self-report and for co-twin ratings, the correlation between measurement errors might be expected to be substantial. Models that constrained the parameters of the reciprocal rating model to be equal across zygosity groups gave a very poor fit to the data, compelling us to relax this constraint both in the analyses of the Neuroticism data and in the analyses of the Extraversion data.

For Neuroticism, we found no evidence for a systematic rater bias (i.e., $f_{RB} = 0$, implying that there is no systematic tendency for respondents high on Neuroticism to overestimate, or underestimate, the Neuroticism of their co-twins). This contrasts with the finding of Silberg et al. (1991), using a questionnaire measure of lifetime history of major depression, who did find a substantial and positive rater bias effect, namely, a tendency for individuals with a history of depression to be more likely to assume a history of depression in their co-twin. The strong rater bias effect reported by Silberg et al. may reflect the greater inaccuracy of ratings of state phenomena (e.g., episodes of major depression) than of ratings of stable personality traits. Neale and Stevenson (1989) reported a slight negative rater bias effect for Emotionality, a construct closely related to Neuroticism, in their analysis of self-report and spouse ratings by the parents of juvenile twins, using the Emotionality, Activity, Sociability, and Impulsivity (EASI) temperament scales, but the statistical significance of the effect was marginal.

For Neuroticism, we did find evidence for a significant correlation between measurement errors for self-report versus co-twin rating data, a correlation that was much higher for monozygotic pairs than for dizygotic pairs (.49 versus .11), perhaps reflecting a much more pronounced tendency for monozygotic pairs, in cases of uncertainty about their co-twin's feelings, to check the same response options for their co-twins as for themselves. The parameter estimates under the best fitting reciprocal rating model (see Figure 2) indicated that residual (i.e., error) variances for co-twin ratings were substantial (55%) but similar in magnitude to the residual variance for self-reports (49%) and that the self-report–true score and co-twin rating–true score correlations were comparable (.71 versus .67), implying that ratings by the co-twin were almost as accurate as self-reports. It should be noted, however, that these estimates are model dependent, so that alternative reciprocal rating models, used with other experimental designs, might lead to different conclusions.

For Extraversion, we did find evidence for a significant but modest rater bias, with extraverted twins tending to underestimate, and introverted twins tending to overestimate, the Extraversion of their co-twin. Although Neale and Stevenson (1989) also found evidence for rater bias for a related construct, Sociability, in their spousal data, the bias was positive rather than negative. We found that monozygotic twin pairs were able to provide more accurate ratings of their co-twins than were dizygotic pairs, although the difference in accuracy was small. Correlations between ratings by the co-twin and the twin's true score (derived from the parameter estimates under the best fitting reciprocal rating model, summarized in Figure 3) were .78 for monozygotic pairs and .69 for dizygotic pairs, (i.e., not much smaller than the self-report–true score correlations of .79 and .82, respectively). Error variances for co-twin ratings were correspondingly larger for dizygotic pairs (51%) than for monozygotic pairs (38%), although error variances for self-report data were comparable (38% and 33%, respectively). In contrast to Neuroticism, for Extraversion, correlations between measurement errors for self-report versus co-twin ratings were modest (.14) and did not vary as a function of zygosity.

Generalizability of Findings From Twin Data

Inferences about the relative importance of genetic and environmental influences on personality in this study depend on our ability to generalize from twin data to other relationships. Unrepresentative sampling of twin pairs would seriously limit our ability to make successful predictions for new relationships. To determine whether there had been deviations from random sampling, we tested for differences in mean and variance of both self-report and co-twin rating personality scores as a function of zygosity and as a function of whether the respondent was from a twin pair concordant for participation in the study or whether the co-twin had refused to cooperate. Most plausible forms of nonrandom sampling of twin pairs with respect to the traits under study will lead to such significant differences in mean or variance (Heath et al., 1989; Lykken et al., 1987, Neale, Eaves, et al., 1989). Differences were mostly small or nonsignificant, except that a larger variance in Extraversion scores was observed for dizygotic than for monozygotic twin pairs, a difference that is explained by the finding of reciprocal inhibitory environmental interaction for Extraversion. On the basis of these results, it is unlikely that our estimates of genetic and environmental parameters will have been seriously biased by deviations from random sampling.

Genetic analysis of data on twin pairs reared together depends on the important assumption that the trait-relevant environments of monozygotic pairs are no more highly correlated than the trait-relevant environments of dizygotic pairs (e.g. Heath et al., 1989). Because there are likely to be many environmental influences of small effect, all contributing additively to personality differences, there is no direct way that we can test this assumption using only data from the present study. The most important test of these assumptions will therefore come from the ability of twin data to predict findings for other relationships. We know of no other genetically informative designs (e.g., adoption data), using adequately large sample sizes, in which self-report personality data have been supplemented with reciprocal rating data on other family members. For self-report data, however, personality data on twins reared together have been found to be in good agreement with adoption, separated-twin, and other family data, once allowance is made for the fact that much of the genetic contribution to Extraversion is nonadditive (e.g., Eaves et al., 1989; Eaves et al., 1992; Loehlin et al., 1985). Consistency of our findings with those of conventional genetic analyses of self-report data, therefore, would imply that any violation of the traditional assumptions of the twin method are having at most a minor effect on the results of our analyses.

Inheritance of Neuroticism

Model fitting confirmed a significant genetic influence on Neuroticism. A nongenetic model gave a very poor fit to the data, even when the correlation between measurement errors for self-report versus co-twin rating was allowed to depend on zygosity. Models that constrained reciprocal rating parameters to be equal across the two zygosity groups suggested a significant contribution of both additive genetic and nonadditive (either dominance or epistatic) genetic effects to personality differences, but they gave a poor fit to the data. Once this constraint on reciprocal rating parameters was relaxed, the evidence for significant genetic nonadditivity disappeared. Under the best fitting model, additive genetic effects accounted for 63% of the variance in trait Neuroticism, with nonshared environmental influences accounting for the remaining 37%. The conclusion that genetic effects on Neuroticism in female respondents are largely additive is in good agreement with the report of Eaves et al. (1992). From an analysis of two large samples of adult twin pairs from Australia (Martin & Jardine, 1986) and Finland (Rose & Kaprio, 1988; Rose et al., 1988), together with new data from a sample of adult American twin pairs and their parents, spouses, siblings, and adult children, Eaves et al. concluded that approximately 75% of the genetic variance in female Neuroticism scores was additive (with the remaining 25% being attributable to epistasis, i.e., to interactions between genes). The estimate of the proportion of the variation in Neuroticism attributable to genetic effects was lower in the Eaves et al. analyses (51.3% for Neuroticism in women) than in the present study. This difference may be explained, however, by the fact that Eaves et al. did not correct their heritability estimate for measurement error, whereas the present analysis, by including both self-report and informant rating data, explicitly allows for measurement error effects.

Inheritance of Extraversion

Analyses of the Extraversion data also confirmed a significant genetic contribution to personality differences. Nongenetic models were rejected by chi-square test of goodness of fit, even when the reciprocal-rating component of the model allowed for zygosity-dependent residual parameters and zygosity differences in sibling knowledge. From the parameter estimates under the best fitting model (see Figure 3), allowing for both the direct genetic effects on Extraversion and the indirect effects arising through the inhibitory environmental effect of co-twin's Extraversion (as discussed in the Method section), we can compute that genetic factors account for 73% of the variance in true Extraversion, and the direct and indirect effects of nonshared environment account for the remaining 27% of the variance.

Although our finding of a substantial additive genetic contribution to variation in Extraversion is consistent with analyses of other major twin data sets (Eaves et al., 1989; Eaves et al., 1992), our results differ in finding a substantial effect of reciprocal inhibitory sibling environmental interaction for Extraversion, namely, a tendency for Extraversion in one twin to inhibit Extraversion in the co-twin, and vice versa. Other analyses have suggested instead a substantial nonadditive genetic contribution to differences in Extraversion (Eaves et al., 1989; Eaves et al., 1992), most probably arising through interactions between genes (epistasis: Eaves et al., 1992). It is possible that nonadditive and additive genetic effects and reciprocal sibling interaction are all contributing to individual differences in Extraversion. A sibling interaction effect for Extraversion has been considered a possibility in several analyses of self-report twin data for Extraversion (Eaves et al., 1989; Jinks & Fulker, 1970). The statistical power for resolving genetic nonadditivity and inhibitory sibling interaction against a background of additive genetic effects on personality is rather poor for twin data (Jardine, 1985), so our failure to find significant nonadditive genetic effects in addition to significant sibling interaction in the present analyses is not surprising. However, it is also possible that the format of our reciprocal rating questionnaire, in which self-report and co-twin rating columns were juxtaposed, encouraged respondents to contrast their personalities with those of their co-twin and thus created or exaggerated a sibling interaction effect. It is less easy, however, to explain why this phenomenon occurred only for Extraversion. Replication of our findings using a modified assessment procedure, in which co-twin ratings and self-report data are obtained independently, would be desirable. Such a procedure might also reduce the strongly correlated errors for self-reports and for co-twin ratings that were observed in monozygotic twin pairs for Neuroticism.

Utility of Reciprocal Rating Data

Our analyses suggest that reciprocal rating data can provide a useful supplement to conventional self-report measures (see also Neale & Stevenson, 1989). In genetic studies, by combining self-report data with reciprocal rating data on family members, it is possible to allow explicitly, by formulation of an appropriate structural model, for some of the biases in self-perception and in perceptions of others, which would otherwise cloud interpretation of the data. In addition to providing evidence for genetic influences on personality that does not rely entirely on the self-perceptions of respondents, reciprocal rating data have several attractions. Measures of Extraversion and Neuroticism are known to be somewhat sensitive to state effects when used with clinical (e.g., depressed) populations (e.g., Coppen & Metcalfe, 1965; Dodwell, 1988; Hirschfeld et al., 1983; Katz & McGuffin, 1987; Knowles & Kreitman, 1965). Other systematic response biases, for example, a tendency to "fake good," are known to influence self-report assessments, particularly of Neuroticism (e.g., Furnham, 1986). In such cases, data from informants may provide a better estimate of stable personality traits than self-report data. In studies in which some self-report data are unavailable, through noncooperation or unavailability of some subjects (e.g., in family studies of the inheritance of personality), the existence of rating data on noncooperative respondents can provide an important check on whether data are missing at random.

For these diverse purposes, our results, which show good validity of simple informant rating measures of Extraversion and Neuroticism when compared with respondent's self-reports, are encouraging. It is important, however, to recognize some of the limitations of the specific application of the reciprocal rating model presented in this article. The reciprocal rating model allowed for rater bias to be a function of the personality of the rater, but it is clearly possible that other variables (e.g., socioeconomic status) may influence both ratings of others and self-reports. Our data consisted solely of reciprocal ratings of twin pairs, but reciprocal ratings on three or more relatives would be needed to determine whether rater bias effects are specific to particular relationships or reflect a generalized tendency for the personality of the rater to influence his or her ratings of others. A simple extension of the reciprocal rating model to allow for ratings of multiple individuals and for self-reports and reciprocal ratings on multiple variables would considerably extend our ability to control for biases in self-perceptions and in perceptions of others.

References

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistences in behavior. *Psychological Review, 81,* 506–520.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science, 250,* 223–228.

Bulmer, M. G. (1980). *The mathematical theory of quantitative genetics.* Oxford, England: Clarendon Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics, 16,* 319–342.

Cattell, R. B. (1957). *Personality and motivation structure and measurement.* London: Harrap.

Condon, J. T. (1988). The assessment of Type A behavior pattern: Results from a spouse-report approach. *Psychological Medicine, 18,* 747–755.

Coppen, A., & Metcalfe, M. (1965). Effect of a depressive illness on MPI scores. *British Journal of Psychiatry, 111,* 236–239.

Crandall, R. (1976). Validation of self-report measures using ratings by others. *Sociological Methods and Research, 4,* 380–400.

Dodwell, D. (1988). Comparison of self-ratings with informant-ratings of premorbid personality on two personality rating scales. *Psychological Medicine, 18,* 495–501.

Eaves, L. J. (1976). A model for sibling effect in man. *Heredity, 36,* 205–214.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture and personality: An empirical approach.* San Diego, CA: Academic Press.

Eaves, L. J., Heath, A. C., Neale, M. C., Hewitt, J. K., & Martin, N. G. (1992). *Sex differences and non-additivity in the effects of genes on personality.* Manuscript submitted for publication.

Eaves, L. J., & Last, K. A. (1980). Assessing empathy in twins through their mutual perception of social attitudes. *Personality and Individual Differences, 1,* 174–176.

Eysenck, H. J. (1960). *The structure of human personality.* London: Routledge & Kegan Paul.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire.* London: Hodder & Stoughton.

Eysenck, H. J., Eysenck, S. B. G., & Barrett, P. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences, 6,* 21–29.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences, 7,* 385–400.

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Sundet, J., Nance, W. E., Corey, L., & Magnus, P. (1985). No decline in assortative mating for educational level. *Behavior Genetics, 15,* 349–369.

Heath, A. C., Jardine, R., Eaves, L. J., & Martin, N. G. (1988). The genetic structure of personality: I. Phenotypic factor structure of the EPQ in an Australian sample. *Personality and Individual Differences, 9,* 59–67.

Heath, A. C., Kendler, K. S., Eaves, L. J., & Markell, D. (1985). The resolution of cultural and biological inheritance: Informativeness of different relationships. *Behavior Genetics, 15,* 439–465.

Heath, A. C., Neale, M. C., Hewitt, J. K., Eaves, L. J., & Fulker, D. W. (1989). Testing structural equation models for twin data using LISREL. *Behavior Genetics, 19,* 9–35.

Hirschfeld, R. M. A., Klerman, G. L., Clayton, P. J., Keller, M. B., MacDonald-Scott, P., & Larkin, B. (1983). Assessing personality effects of the depressive state on trait measurement. *American Journal of Psychiatry, 140,* 695–699.

Jardine, R. (1985). *A twin study of personality, social attitudes and drinking behavior.* Unpublished doctoral dissertation, Australian National University, Canberra, Australian Capital Territory.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical, genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin, 73,* 311–349.

Jöreskog, K. G., & Sörbom, D. (1983). *LISREL: Analysis of linear structural relationships by the method of maximum-likelihood.* Chicago: International Educational Resources.

Jöreskog, K. G., & Sörbom, D. (1986). *PRELIS: A preprocessor for LISREL.* Mooresville, IN: Scientific Software.

Kasriel, J., & Eaves, L. J. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science, 8,* 263–266.

Katz, R., & McGuffin, P. (1987). Neuroticism in familial depression. *Psychological Medicine, 17,* 155–161.

Kendler, K. S., Kessler, R. C., Heath, A. C., Neale, M. C., & Eaves, L. J. (in press). Coping: A genetic epidemiologic investigation. *Psychological Medicine.*

Knowles, J. B., & Kreitman, N. (1965). The Eysenck Personality Inventory: Some considerations. *British Journal of Psychiatry, 111,* 755–759.

Loehlin, J. C. (1981). Personality resemblances in adoptive families. *Behavior Genetics, 11,* 309–330.

Loehlin, J. C., & Nichols, R. C. (1976). *Heredity, environment and personality: A study of 850 sets of twins.* Austin, TX: University of Texas Press.

Loehlin, J. C., Willerman, L., & Horn, J. (1985). Personality resemblances in adoptive families when the children are late-adolescent or adult. *Journal of Personality and Social Psychology, 48,* 376–392.

Lykken, D. T., McGue, M., & Tellegen, A. (1987). Recruitment bias in twin research: The rule of two-thirds reconsidered. *Behavior Genetics, 17,* 343–362.

Martin, N. G., Eaves, L. J., Kearsey, M. J., & Davies, P. (1978). The power of the classical twin study. *Heredity, 40,* 97–116.

Martin, N. G., & Jardine, R. (1986). Eysenck's contribution to behavior genetics. In S. Modgil & C. Modgil (Eds.), *Hans Eysenck: Consensus and controversy* (pp. 213–218). Sussex, England: Falmer Press.

Martin, N. G., & Wilson, S. R. (1982). Bias in the estimation of heritability from truncated samples of twins. *Behavior Genetics, 12,* 467–472.

McGue, M., Wette, R., & Rao, D. C. (1989). Path analysis under generalized marital resemblance: Evaluation of the assumptions underlying the mixed homogamy model by the Monte Carlo method. *Genetic Epidemiology, 6,* 373–388.

Neale, M. C., Eaves, L. J., Hewitt, J. K., & Kendler, K. S. (1989). Bias in correlations from truncated samples of relatives. *Behavior Genetics, 19,* 163–169.

Neale, M. C., Heath, A. C., Hewitt, J. K., Eaves, L. J., & Fulker, D. W. (1989). Fitting genetic models with LISREL: Hypothesis-testing. *Behavior Genetics, 19,* 37–49.

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI Temperament scales: A twin study. *Journal of Personality and Social Psychology, 56,* 446–455.

Pedersen, N. L., Plomin, R., McClearn, G. E., & Friberg, L. (1988). Neuroticism, extraversion, and related traits in adult twins reared apart and reared together. *Journal of Personality and Social Psychology, 55,* 950–957.

Rose, R. J., & Kaprio, J. (1988). Frequency of social contact and intrapair resemblance of adult monozygotic cotwins—Or does shared experience influence personality after all? *Behavior Genetics, 18,* 309–328.

Rose, R. J., Koskenvuo, M., Kaprio, J., Sarna, S., & Langinvainio, H. (1988). Shared genes, shared experiences, and similarity of personality: Data from 14,228 Finnish co-twins. *Journal of Personality and Social Psychology, 54,* 161–171.

Scarr, S., Webber, P. L., Weinberg, R. A., & Wittig, M. A. (1981). Personality resemblance among adolescents and their parents in biologically related and adoptive families. *Journal of Personality and Social Psychology, 40,* 885–898.

Shields, J. (1962). *Monozygotic twins: Brought up apart and brought up together.* London: Oxford University Press.

Silberg, J. L., Meyer, J. M., Eaves, L. J., Neale, M. C., Hewitt, J. K., Heath, A. C., & Kendler, K. S. (1991). The effect of bias on the estimates of the genetic and environmental influences on ratings of depressive symptoms. *International Journal of Methods in Psychiatric Research, 1,* 59–67.

Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K. J., Segal, N., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology, 54,* 1031–1039.

Received April 1, 1991
Revision received December 23, 1991
Accepted December 26, 1991 ■