JOURNAL OF RESEARCH IN PERSONALITY **32,** 431–453 (1998)
ARTICLE NO. RP982225

Heritabilities of Common and Measure-Specific Components of the Big Five Personality Factors

John C. Loehlin

University of Texas at Austin

Robert R. McCrae and Paul T. Costa, Jr.

Gerontology Research Center, National Institute on Aging, National Institutes of Health

and

Oliver P. John

University of California, Berkeley

Three different measures of the Big Five personality dimensions were developed from the battery of questionnaires used in the National Merit Twin Study: one from trait self-rating scales, one from personality inventory items, and one from an adjective check list. Behavior-genetic models were fit to what the three measures had in common, and to the variance distinctive to each. The results of the model fitting agreed with other recent studies in showing the Big Five dimensions to be substantially and about equally heritable, with little or no contribution of shared family environment. Heritabilities for males and females did not differ significantly. For Agreeableness and Conscientiousness, some effect of shared environment was found for measure-specific variance on the personality inventory, and for Extraversion and Neuroticism, models involving nonadditive genetic variance or twin contrast effects provided slightly better fits.

Key Words: Big Five; heredity; environment; twins; model fitting; sex differences.

In recent years, there has been considerable interest in the so-called ''Big Five'' dimensions of personality, which were originally derived from the vast array of trait terms used in natural languages, and which appear to have

We thank Lewis R. Goldberg and Robert C. Nichols for their helpful comments on an earlier version of the manuscript.

Correspondence and reprint requests should be addressed to J. C. Loehlin, Department of Psychology, University of Texas, Austin, TX 78712. E-mail: loehlin@psy.utexas.edu.

0092-6566/98 $25.00

considerable generality across measuring instruments, languages, and methods of analysis (for reviews, see Digman, 1990; John, 1990; and McCrae, 1989). The names given to these dimensions sometimes differ: we will refer to them as Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N), and Openness (O).

Although not all psychologists are enamored of the Big Five (e.g., Block, 1995), and many others prefer to work at different levels of detail or with other types of constructs than traits, there has been sufficient interest in the Big Five to justify looking at genetic and environmental contributions to their variation. There have been several recent assessments of the extent to which genes or family environments account for individual differences on the Big Five dimensions. Most of these come to the conclusion that all five are moderately and about equally heritable, and that shared family environmental factors play a minor role, if any, in accounting for individual differences on them.

Loehlin (1992) summarized the then-existing evidence from twin, adoption, and family studies for a variety of questionnaire scales that he classified according to the Big Five. He fit behavior-genetic models to the available correlations from the literature, and concluded that under various assumptions the genes contributed from 28% to 59% of the variance of the Big Five traits, and the shared environment of family members 0% to 11%. Among the studies included in the model fitting, the only one that included explicit measures of A,C, and O, a Swedish study of elderly twins, produced typical results for C and O, but a different result for A. The genetic contribution to measured Agreeableness in this study, estimated at 12%, did not differ significantly from zero, and the family environment contribution, estimated at 21%, was statistically significant (Bergeman et al., 1993).

A recent heritability analysis of the Big Five dimensions as measured by the Revised NEO Personality Inventory of Costa and McCrae (1992) used a total of 660 monozygotic (MZ) pairs and 380 dizygotic (DZ) pairs from pooled Canadian and German twin samples (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998). For all five traits a simple model involving only additive genes and nonshared environment fit the data. Estimates for the heritabilities of factor scales for E, A, C, N, and O were .50, .48, .49, .49, and .48, respectively. These estimates are obviously very similar—in particular, the heritability of Agreeableness is quite in line with the rest.

Another recent twin study, this time in California, measured 313 MZ pairs and 91 DZ pairs from a volunteer community sample (Waller, in press). Waller used seven factor scales, five of them corresponding fairly closely to the Big Five. The heritabilities were .49, .33, .46, .42, and .58, for scales corresponding to E, A, C, N, and O. Again, Agreeableness shows clear evidence of heritability, although it is a little lower than the others. Evidence for shared family environmental effects was not found for Agreeableness, being present only for Neuroticism, and there not large (12%).

In the German study, ratings were obtained from two peers of the twins as well as from the twins themselves. The authors carried out a joint model fitting to the self- and peer-report data (Riemann, Angleitner, & Strelau, 1997). From this, the estimates of the genetic contribution to the Big Five dimensions were .60, .57, .71, .61, and .81, respectively, with the effects of shared environment again being negligible. The estimates of genetic effects are substantial for each of the Big Five dimensions. Furthermore, the heritabilities were higher than those typical for twin self-report data alone. Apparently, the common element distilled from various perspectives on a given individual was more heritable than that unique to a particular viewpoint.

Global and Specific Levels of Analysis

Most models of personality are hierarchical, with broad factors (such as the Big Five) defined by more narrow and specific traits. Heritability can be studied at either level. Loehlin (1992) fit data for two components of Extraversion—Dominance and Sociability—from several twin, family, and adoption studies. Similar estimates were obtained for the additive effects of genes (36% in each case) but different estimates for nonadditive genes (22% for Dominance, 8% for Sociability).

In the combined German-Canadian study, the authors were interested in examining what was distinctive to various facets of each Big Five dimension, in addition to what the facets had in common (Jang, et al., 1998). To this end the authors fit models to the residual facet scores in each domain, after regressing out the common factors. For the most part, these residual scores still fit a genes-and-unshared-environment model, but with lower heritabilities—typically in the .20s and .30s. One of the reasons for the lower heritability estimates is no doubt the relatively large proportion of variance in these residual scores that represents measurement error. However, the residuals from a couple of facets in the Agreeableness domain, Altruism and Modesty, appeared to fit a shared environment model, as did Achievement Striving and Deliberation from the Conscientiousness domain.

The present study looks at archival data from a large U.S. twin sample from the perspective of the Big Five. These were data that had been gathered prior to the widespread use of this approach to describing personality. The sample consisted of the 839 twin pairs of the National Merit twin study, as described in Loehlin and Nichols (1976). Thus this study allows us to extend heritability estimates for the Big Five to a somewhat different population and age group: college-bound U.S. high school seniors in the early 1960s. Moreover, it proved possible to estimate the Big Five dimensions separately from three different sources within this data set: (1) from a series of 47 bipolar trait self-ratings, (2) from the items of a standard personality questionnaire, the *California Psychological Inventory* (CPI, Gough, 1957), and (3) from a version of Gough's *Adjective Check List* (ACL, Gough & Heilbrun, 1965). Thus we are in position to examine for each trait both the heritability of the factor common to these different response modes, and the heritabilities of the distinctive contributions of each mode. Although all three response modes (trait ratings, CPI, and ACL) were designed to measure the five factors at a global level, they differ somewhat in the specific component traits they emphasize, so the analysis of specific variance in each mode is in some respects like Jang et al.'s analysis of specific facet variance. The three modes also differ in their approach to measurement, and in that respect our design resembles the analysis of variance common and distinctive to self- and peer-ratings. Most basically, these analyses address the question of whether different indicators of a global trait are interchangeable from a behavior genetics viewpoint, or whether results are likely to vary depending on the indicator used.

METHOD

Participants

The sample was drawn from among the roughly 600,000 high school juniors who took the National Merit Scholarship Qualifying Test in 1962. An item on the face sheet asked the test-taker if he or she was a twin. Pairs of individuals answering ''yes'' to this question who were from the same high school, of the same sex, and had the same surname and home address were matched by computer. They were then sent a brief questionnaire confirming that they were indeed twins and inviting them to participate further in the study, as well as asking questions about physical resemblance, mistaken identity, and the like, which could be used to diagnose the pair as MZ or DZ. Those agreeing to participate were individually sent a battery of questionnaires assessing their personalities, activities, interests, and values, which they returned by mail. A parent of the twins responded to a questionnaire about the twins' early behavior and experiences. Further details and copies of the questionnaires may be found in Loehlin and Nichols (1976). Most results to be reported are from 807 pairs of twins, 490 diagnosed as MZ and 317 as DZ.

Measures

Trait ratings. The questionnaire filled out by the twins contained a set of 47 self-rating scales. Each was a 7-point scale identified by a word or phrase at each end, such as ''Messy–Neat,'' ''Popular–Unpopular,'' ''Religious–Nonreligious.'' The testee was instructed to circle one of the seven numbers along the scale, which were successively labeled ''Very,'' ''Fairly,'' and ''Slightly'' at each end, with ''Neither or Both'' in the middle.

Four expert judges[1] experienced with the Big Five classified the 47 scales into up to two categories each (including a residual category ''Beyond the Big Five''). Scales were developed based on these classifications. In most cases, there was unanimous agreement among the four

[1] The judges, in addition to the fourth author, were Lewis R. Goldberg, John M. Digman, and Maureen Barckley. We are grateful for their expert assistance.

TABLE 1
Big Five Trait-Ratings Scales, with Sample Items

Extraversion (5 items; α = .75)	
	Shy–Outgoing[a]
	Timid–Bold[a]
Agreeableness (6 items; α = .63)	
	Easily angered–Good-natured[a]
	Kind[a]–Cruel
Conscientiousness (9 items; α = .80)	
	Responsible[a]–Irresponsible
	Dependable[a]–Undependable
Neuroticism (7 items; α = .68)	
	Happy–Unhappy[a]
	Well-adjusted–Maladjusted[a]
Openness (3 items; α = .20)	
	Original[a]–Unoriginal
	Sophisticated[a]–Unsophisticated

[a] Direction of response scored for the trait.

judges on the classification of the ratings used for each scale. Ratings on the subscales were individually standardized and summed to yield Big Five scale scores.

The scales are described in Table 1. (The sample items given in this and the following tables are the two having the highest item-total correlations.) Considering their relative brevity, four of the five scales have reasonable alpha reliabilities. The fifth, Openness, had very few rating items judged to represent it, and its low α of .20 in part reflects the shortness of this three-item scale.

CPI scales. Included with questionnaire battery sent to the National Merit twins was a standard personality inventory, the *California Psychological Inventory* (CPI). Big Five scales were developed for the CPI based on a sample of 348 adults from the Baltimore Longitudinal Study of Aging (BLSA; Shock et al. 1984) who had been given both the NEO-PI and the CPI. The sample included 153 men aged 27 to 92, and 195 women aged 19 to 89 (McCrae, Costa, & Piedmont, 1993).

For each of the five factors, each CPI item was first correlated with the NEO-PI criterion factor within gender. Only items that were significantly correlated ($p < .05$) with the factor in both groups were retained. In the next step, these items were correlated with all five factors in the combined sample, and items that correlated more highly with one of the other factors were eliminated, to yield nonoverlapping CPI scales. These final scales contained between 16 and 75 items, and had α reliabilities between .72 and .93, as shown in Table 2. As might be expected, the α reliabilities tend to vary inversely with scale length, with the lowest occurring for the 16-item scale for Agreeableness.

Some evidence on the reasonableness of these Big Five CPI scales is given by correlations with other personality measures collected in the BLSA, such as the *Myers–Briggs Type Indicator* (MBTI, Myers & McCaulley, 1985) and the Jackson (1984) *Personality Research Form* (PRF). Correlations with the MBTI corresponded very closely both in pattern and size of correlations to those obtained in an earlier study using the NEO-PI (McCrae & Costa, 1989). Correlations with the PRF scales were also generally similar to those obtained earlier with the NEO-PI (Costa & McCrae, 1988), although those involving the A and C scales tended to be somewhat smaller than the correlations obtained in that study. Further details on the sam-

TABLE 2
Big Five CPI Scales, with Sample Items

Extraversion (51 items; α = .92)
I am a good mixer (T)
It is hard for me to find anything to talk about when I meet a new person (F)
Agreeableness (16 items; α = .72)
I have one or more bad habits that are so strong that it is no use fighting against them (F)
I am often said to be hotheaded (F)
Conscientiousness (29 items; α = .82)
I set a high standard for myself and I feel others should do the same (T)
I always try to do at least a little better than what is expected of me (T)
Neuroticism (75 items; α = .93)
I certainly feel useless at times (T)
I often get disgusted with myself (T)
Openness (42 items; α = .83)
Disobedience to any government is never justified (F)
I read at least ten books a year (T)

Note. Keyed response in parentheses after each item. CPI = California Psychological Inventory.

ples, scale construction, and correlations with other measures are available from the second author.

Adjective scales. A 159-item abbreviated version of Gough and Heilbrun's (1965) *Adjective Check List* (ACL) was part of the battery of questionnaires filled out by the twins. In a study by John and Roberts (1993), adjectives from the full ACL were classified according to the Big Five dimensions. From the ACL adjectives identified by John and Roberts as belonging to one of the Big Five factors, those included in the version filled out by the twins were taken to define the Big Five adjective scales. Further procedural details may be obtained from the fourth author.

The tendency of an individual to mark many or few adjectives is a pervasive artifact in adjective check list studies. Therefore, total endorsement frequency, the number of items out of the 159 that each individual endorsed as self-descriptive, was removed from the raw scores by regression.

The properties of the Big Five adjective checklist scales are shown in Table 3. Except for A, the scale lengths lie between those for the CPI scales and the trait self-rating scales, and for the most part the alpha reliabilities are intermediate as well. Note that there is some direct content overlap between the trait-rating and the adjective scales. Several of the end-point labels of the trait scales, for example, Outgoing, Responsible, and Original, also appear as adjectives scored on the checklist scales.

Intelligence and Openness. There has sometimes been discussion of the fifth of the Big Five as a possible manifestation of intelligence in the personality domain. In our sample there was a substantial correlation between the National Merit Scholarship Qualifying Test (NMSQT) score and Openness as measured by the CPI items (r = .50). Correlations of the NMSQT with the other two Openness scales were also positive, but lower (.22 and .08, for adjectives and self-ratings, respectively). Because we wished to emphasize the status of Openness as a *personality* dimension, overall intellectual ability, as measured by NMSQT total score, was removed from the three Openness scores by regression.

TABLE 3
Big Five Adjective Checklist Scales, with Sample Items

Extraversion (21 items; α = .72)
Outgoing (T)
Talkative (T)
Agreeableness (16 items; α = .75)
Helpful (T)
Pleasant (T)
Conscientiousness (23 items; α = .81)
Responsible (T)
Efficient (T)
Neuroticism (20 items; α = .81)
Worrying (T)
Tense (T)
Openness (12 items; α = .65)
Original (T)
Imaginative (T)

Note. Keyed response in parentheses after each item.

It is recognized that something may be lost by doing this—the observed personality trait of Openness has legitimate intellectual aspects, and in some sense one might be removing too much by regressing intelligence out. Nevertheless, for our purposes the gain in clarity justifies this step. Much is known concerning the heritability of intelligence, but less concerning the personality side of Openness.

RESULTS

Simple Correlations

Table 4 gives MZ and DZ twin correlations for the Big Five measures from the three data sources. It will be noted that the MZ correlations are uniformly higher than the DZ correlations, suggesting a genetic contribution

TABLE 4
MZ and DZ Twin Correlations for the Three Sets of Big Five Scale Scores

	Ratings		CPI		ACL	
Factor	MZ	DZ	MZ	DZ	MZ	DZ
Extraversion	.47	.01	.60	.30	.39	−.06
Agreeableness	.32	.06	.46	.34	.29	.18
Conscientiousness	.42	.21	.53	.34	.37	.14
Neuroticism	.43	.17	.53	.25	.44	.06
Openness	.39	.19	.49	.27	.36	.08
Mean	.41	.13	.52	.30	.37	.08

to these measures. For the CPI scales, the DZ correlations are typically about half as large as the MZ correlations, consistent with a model of additive genes and little effect of shared family environment—the common finding in earlier studies. The DZ correlations for the trait self-rating scales and the adjective check list are often considerably less than half the MZ correlations, suggesting the possible presence of some additional factor such as nonadditive genetic variance or contrast effects. However, the measures in Table 4 contain both common and unique factors, and the model-fitting to be undertaken will show more clearly whether the data are consistent with simple models or require more complex ones.

Factor Analysis of Measures

To verify that the 15 measures did indeed define five factors corresponding to the Big Five, a preliminary factor analysis was carried out, based on the 1614 individuals in the sample. Five principal factors were extracted from a correlation matrix, using squared multiple correlations as communality estimates, and rotated by Varimax. The resulting factor loadings are shown in Table 5.

On the whole, the factor loadings correspond to those expected for Big Five factors (boldface type). There are a few fairly weak defining loadings, but all are over .30; for every factor, at least two of the three exceed .50.

TABLE 5
Factor Analysis of the 15 Measures

	Factor				
Measure	1(E)	2(A)	3(C)	4(N)	5(O)
Rating-E	**.90**	−.02	.06	−.02	.14
Rating-A	−.00	**.85**	.16	−.06	.01
Rating-C	.23	.26	**.81**	.10	.09
Rating-N	−.29	−.47	−.22	**.35**	−.04
Rating-O	.15	−.04	.11	.03	**.62**
CPI-E	**.73**	−.00	.11	−.08	.21
CPI-A	−.20	**.32**	.13	−.34	−.08
CPI-C	−.04	.06	**.55**	−.21	−.06
CPI-N	−.12	−.21	−.24	**.71**	−.16
CPI-O	.08	−.05	−.11	−.06	**.41**
ACL-E	**.75**	.04	−.02	−.04	−.03
ACL-A	−.05	**.60**	.07	−.21	−.16
ACL-C	.01	.11	**.65**	−.27	−.01
ACL-N	−.20	−.45	−.17	**.53**	−.10
ACL-O	−.01	.00	.02	−.05	**.59**

Note. Five principal factors extracted from correlation matrix with squared multiple correlations as communalities, and rotated via Varimax. Identifying loadings in boldface.

There are three off-pattern loadings of .30 or more, all involving a negative relationship between A and N. This probably reflects the presence of an evaluative component, as A tends to be the most positively evaluated of the Big Five, and N the most negative.

Model Fitting

It is usual in behavior-genetic analyses of complex multivariate data to fit explicit models to covariance matrices among the variables involved. This permits one to ask whether the data are consistent with simple models, or whether more complex models are required. If models are found that do fit the data, estimates of their parameters are available, from which estimates of the relative contributions of genes and environments to the variation of the trait may be obtained (Neale & Cardon, 1992).

The model fitting was based on three measures for each of the Big Five dimensions. One of these measures was from the trait ratings, one from the CPI, and one from the ACL, obtained as described in the preceding section. All three measures were standardized over the total set of individuals in the sample, and covariance matrices calculated separately for MZ and DZ pairs, using these standardized scores. (Had the scores been separately standardized for the MZs and DZs and for twins 1 and 2 these would be correlation matrices, but we did not do this in order to retain the statistical advantages of covariance matrices for model fitting.)

The model fitting was done separately for each of the Big Five dimensions because they are in theory independent. The data to be fit in each case consist of two six-variable covariance matrices, one for MZ twins and one for DZ twins, containing the three measures for the first twin of each pair followed by the same three measures for the second twin. (The covariance matrices are given in the Appendix.) As a first step, a traditional twin model involving additive genes (*h*), shared or common environment (*c*), and unshared environment (*e*), was fit to what the three scales measuring the factor had in common, and to the residual variance distinctive to each type of scale.

A path diagram of the model is given in Fig. 1. One twin, indicated by subscript 1, is represented in the lefthand part of the figure, the other, by subscript 2, on the right. The top part of the figure represents the genetic and environmental causation of the latent Big Five factor (labeled F). The factor is a latent trait indexed by the three measures R (for ratings), Q (for questionnaire items), and A (for adjectives). At the top of the figure are the genotypes, G, connected to F via a path *h*, and correlated across twins 1.0 for MZs and .5 for DZs. (The value of .5 for DZs assumes that the genetic variance in G is purely additive and the twins' parents mated randomly with respect to this trait. Assortative mating tends to be fairly slight for personality traits [e.g., Price & Vandenberg, 1980]; possible nonadditive genetic effects will be examined later.) Also at the top of the figure are shown the twins'

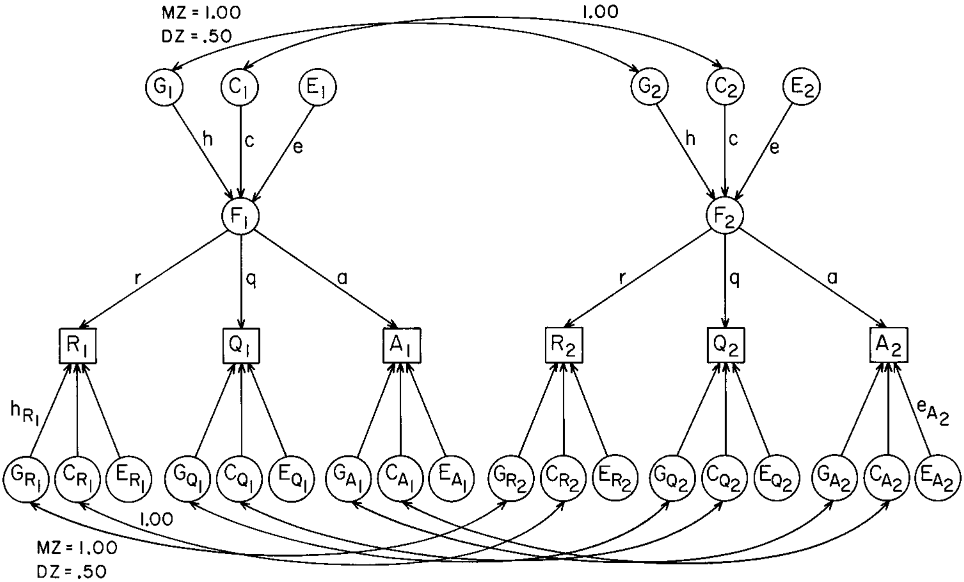

FIG. 1. Path model of common and measure-distinctive genetic and environmental causation of a Big Five factor and scales measuring it. (Circles = latent variables, squares = manifest variables; straight arrows = causes, curved arrows = correlations; 1,2 = the two twins of a pair; F = Big Five factor; G, C, E = additive genes, common environment, unshared environment; R, Q, A = trait ratings, questionnaire items, adjective scales; lower case letters = corresponding causal paths. Only a few paths are labeled in the bottom part of the figure; analogous labels apply for the rest.)

shared environments C, connected to the latent trait by a path *c* and correlated (by definition) 1.0 between the twins. A third path to F, *e*, accounts for the residual variation, E, reflecting environmental events, pre- or postnatal, that are unique to the particular twin, plus genotype-environment interaction—but not measurement error, because F is a latent trait. Effects general to the testing situation but specific to the individual, such as current mood, could also be part of E.

The observed measures are shown in the squares, labeled R for the trait ratings scale, Q for the CPI items scale, and A for the ACL scale. They are connected to the factor via the paths *r*, *q*, and *a*, which represent factor loadings—the extent to which the common factor F contributes to the particular observed scores.

The bottom part of the figure represents the causation of the specific parts of each of the observed measures of the trait, i.e., those aspects independent of the common Big Five factor. Below the squares are Gs, representing the genetic contribution to the specific part; Cs, representing the shared environ-

TABLE 6
Genetic and Environmental Components of the Variance of the Latent Big Five Dimensions, from Fitting the Simple *h, c, e* Model

Dimension	h^2	c^2	e^2	95% CI for h^2
Extraversion	.57	.00	.44	.50 to .64
Agreeableness	.51	.00	.49	.42 to .61
Conscientiousness	.52	.00	.48	.43 to .61
Neuroticism	.58	.00	.42	.51 to .66
Openness	.56	.00	.44	.45 to .68
Mean	.55	.00	.45	

Note. h^2 = genetic variance; c^2 = shared environmental variance; e^2 = nonshared environmental variance. *h, c, e* = corresponding paths. CI = confidence interval. 95% CI is approximate, obtained as $(h \pm 1.965SE_h)^2$, from model with *c* fixed to zero.

mental contribution; and Es, representing unique environmental contributions, genotype-environment interactions, situational factors, and, this time, measurement error (because R, Q, and A are observed scores). Corresponding Gs are again assumed correlated 1.0 or .5 across twins, depending on zygosity, and the Cs correlated 1.0 by definition. The specific factors for the three measures are represented as mutually independent. This is so by definition within twins. There are logically possible scenarios whereby specifics for one measure for one twin might become correlated with specifics for a different measure for the other twin, but our modeling assumes that if any such effects exist, they are trivial enough to be ignored without seriously distorting the analysis.

The model was fit to the five pairs of MZ and DZ covariance matrices using the computer program LISREL (version 8—Jöreskog and Sörbom, 1993). The common factor variance was constrained to equal 1.0, and the variances of G, C, and so on were fixed to 1.0. The variances of the observed variables are close to 1.0 because of the preliminary standardization over the combined groups, so the various estimates may be interpreted as standardized values, to a reasonable approximation.

The common factors. The variance contributions for the five latent factors are shown in Table 6. For genes and shared environment they are the squares of the *h* and *c* path values; the residual variances e^2 were obtained as such in the solution.

As is evident from the table, these values proved to be quite similar across the five factors: They range from .51 to .58 for h^2, from .42 to .49 for e^2, and are uniformly estimated as zero for c^2—the minimum value the model permits. This model, then, suggests that, on average, about 55% of individual personality trait variation in this population is associated with genetic differences, 45% with environmental events unique to each twin, situational factors, or genotype-environment interaction, and none to shared environment. Moreover, these contributions are quite similar across the Big Five factors, including an Openness factor that is independent of measured intelligence. (Out of curiosity, we also fit the model to the unpartialed data for Openness. The result was a somewhat poorer overall fit, a higher h^2, 69%, a lower e^2, 32%, and c^2 still zero.) Because the analysis is being made for latent variables, the unique environmental component will not include random errors of measurement, but, as noted, it could include transient factors such as mood that might influence a given individual differently from his or her twin during the period of filling out the questionnaire battery.

Even with these fairly large twin samples the parameter estimates are not extremely precise. Approximate 95% confidence intervals for h^2 are given in the table, obtained via the standard errors estimated by the fitting program (see table footnote for details). Obviously, these confidence intervals would be consistent with modest differences in heritability among the Big Five factors, although (particularly in conjunction with similar results in other twin studies) they suggest that any very large differences of this kind are unlikely.

Table 7 provides a breakdown for each of the three measures into the variance attributable to the common factor and that to the specifics of the measuring instrument, with the specific part divided into the contributions of the genes, the shared environment, and a residual including unique environment, situational factors, genotype-environment interaction, and error.

The first column of Table 7 gives the squares of the paths from the latent variable to each of the three measures. These indicate the extent to which the variance of each measure reflects the common factor, as opposed to specifics and error. Across the Big Five, the trait rating and adjective scales tend to share more in common and the CPI scales less. This is perhaps not surprising, because of the overlap between the trait ratings and the adjective check list—recall that many of the trait rating scales are anchored by descriptive adjectives, some of which also occur in the ACL. The fact that the loadings of the CPI scales are lower presumably reflects distinctive variance in these scales, not error, because the scales are fairly reliable. The three openness scales tend to have lower path values than the others. In this case, this may well represent a somewhat greater proportion of measurement error, given the rather short scales. (It was not a result of the partialing out of ability variance—the unpartialed version did not differ greatly in this respect, with values of .30, .17, and .53.)

The measure-specific factors. The righthand portion of Table 7 contains the heredity-environment analyses for the specifics. More often than not, there are nonnegligible genetic components to the specific variances—for Conscientiousness, Neuroticism, and Openness for the trait ratings; for all but Agreeableness for the CPI; for Extraversion, Conscientiousness, and

TABLE 7
Genetic and Environmental Components of the Variance Specific to Particular Measures of the Latent Big Five Dimensions, from Fitting the Simple *h, c, e,* Model

Dimension & Measure	General part	Specific part		
		h^2	c^2	e^2
Extraversion				
Ratings	.87	.00	.00	.15
Questionnaire	.53	.15	.04	.24
Adjectives	.57	.08	.03	.36
Agreeableness				
Ratings	.52	.04	.00	.42
Questionnaire	.23	.01	.30	.46
Adjectives	.53	.00	.03	.44
Conscientiousness				
Ratings	.50	.17	.00	.30
Questionnaire	.29	.15	.18	.36
Adjectives	.59	.09	.00	.33
Neuroticism				
Ratings	.52	.15	.00	.34
Questionnaire	.41	.24	.01	.33
Adjectives	.70	.02	.00	.30
Openness				
Ratings	.37	.22	.00	.42
Questionnaire	.14	.36	.04	.45
Adjectives	.42	.10	.00	.48

Note. h^2 = genetic variance; c^2 = shared environmental variance; e^2 = nonshared environmental variance.

Openness for the ACL. (We are considering as nonnegligible those components contributing 5% or more to the total variance of a measure; typically, these are also significant at the .05 level by a differential χ^2 test.) By contrast to the common factors, there are some nontrivial shared environmental factors among the specifics—none for the trait ratings or the ACL, but for Agreeableness and Conscientiousness among the CPI scales. There tend to be substantial e^2s throughout for the specifics; it will be recalled that these now include errors of measurement.

We can test whether the specifics truly differ in composition across the three measures by fitting models constraining the *h*s and *c*s to be equal across measures, and seeing if the goodness-of-fit chi square rises significantly. Indeed it does, in all five cases: the *h*s and *c*s the of the specifics differ across measures.

Sex differences. Could better fits be obtained by fitting the model separately for males and females? Covariance matrices were calculated for the two sexes separately, and the basic *h, c, e* model was fit twice, once with the 14 free parameters of the model allowed to differ between the sexes, and once constrained to be the same. Chi square tests of the difference in fit were not statistically significant for A, C, or O, suggesting that sex differences were not important for these three Big Five dimensions.

The overall chi square tests were significant for E and N, indicating the presence of sex differences. However, in both cases, the parameters of greatest interest—*h, c,* and *e* for the latent trait—could be equated across the two sexes without a significant rise in χ^2. This also proved to be the case for the *h*s, *c*s, and *e*s for the specifics. Thus the significant sex differences appeared to lie in the factor loadings. In the case of E, for *r, q,* and *a*, respectively, these were .98, .72, and .79 for males and .85, .76, and .70 for females (with the rest of the parameters equated across the sexes). For N the corresponding values were .67, .67, and .72 for males and .76, .60, and .88 for females. Obviously, in all cases the common factor is expressed substantially in all three indicators, so no gross differences in interpretation are called for. For males, the ratings seem to carry more weight for E, and for females, the ACL for N. Any more elaborate interpretation of these sex differences should probably await their replication in other samples.

Alternative models. Even though the simple model of Fig. 1 fits fairly well by tests to be discussed in the next section, the presence of lower-bound estimates (zeroes) for the shared environment may be resulting from DZ twin correlations less than one-half MZ twin correlations, a situation in which models that include nonadditive genetic variance or contrasts between twins may provide better fits. Two such alternative models are considered. One introduces a nonadditive genetic parameter (*i*) in lieu of shared environment. This was modeled as epistasis (interactions involving several gene loci) by replacing the Cs at the top of Fig. 1 by Is, and setting the correlation between the epistatic genes (I) of the two twins to 1.0 for MZs and zero for DZs. This represents the hypothesis that these traits may be ''emergenic,'' in the sense of Lykken (1982). Models that represented the nonadditive genetic variation as due to dominance rather than epistasis were also fitted, but are not separately reported. The dominance and epistasis models were pretty much equivalent in terms of overall fits and estimates of broad heritability, although the dominance models tended to place more of the genetic variance in the nonadditive category. For a discussion of some other possible genetic models, see Eaves (1988).

The contrast effects model retained a shared environment C, but added a contrast parameter *b* between the two twins, representing the possibility that each twin might be affected in his judgment of himself (or in his actual development) by the other twin's status on the trait. If this parameter is positive, the effect is one of assimilation, making the twins more alike; if negative, the effect is one of contrast, making the twins less alike. This model isn't quite as straightforward to report as the others, because it implies variances that differ somewhat between MZs and DZs because of covariances involving the additional path. We present results standardized on the basis of the DZs, who are expected to be more like the general population.

The various alternative models were applied only at the level of the general factor, although in principle one could fit such models at the level of the specifics as well. (Readers wishing to explore such alternatives can do so using the covariance matrices given in the Appendix.)

Tests of the model fitting. The results from fitting the three models to the data from each of the five factors are shown in Table 8. Information about overall model fit is given in the three rightmost columns of the table. In the first of these columns is the goodness-of-fit chi square, which allows a test of the null hypothesis that the model fits exactly in the population. Next is the root mean square error of approximation (RMSEA), an index that estimates the goodness of approximation to a true model in the population. One rule of thumb for interpreting the RMSEA is that ''values below .10 represent a good fit, values below .05 a very good fit'' (Steiger, 1989, p. 81). By this criterion, all the model fits are good, and all but those for E are very good. Because the sample sizes are large, we can reject the null hypothesis of an *exact* fit for E or A for any of the models (by a χ^2 test with 27 or 28 *df*, $p < .05$). However, exact fits for C, N, or O cannot be ruled out. The final column of the table gives the upper limit of the 90% confidence interval for RMSEA, which allows us to test the null hypothesis that the model provides a *poor* approximation in the population (i.e., RMSEA $\geq$ .10). Mostly, we can reject this null hypothesis of poor fit. Only in the case of E do values of RMSEA above .10 fall within the 90% confidence interval; for the other four factors we can be 90% confident that the fit in the population is a good one for any of the three models.

Can we distinguish among the fits of the three alternative models? In the case of the *h,c,e,b* model, the difference in χ^2s between it and the *h,c,e* model can be tested directly as a chi square with 1 df. In the case of the *h,i,e* model, the difference from an *h,e* model, which will have the same chi square as the *h,c,e* model because of the zero *c* parameter, can be similarly tested. Such tests indicate that for Extraversion and Neuroticism either of the two alternatives fits significantly better than the original model; this is not true for A, C, or O.

In short, one could justify interpreting any of the three models for any of the five factors, although one might prefer one of the two alternative models for E and N. How do these differ in their results from the simple *h,c,e* model? In general, the *h,i,e* model allows some of the genetic variance to shift from additive (h^2) to nonadditive (i^2)—the bulk of it in the case of E, about half of it for N, and lesser amounts for A, C, and O. The total amount of genetic variance, in the column h_b^2 (the so-called ''broad heritability'') stays about the same.

TABLE 8
Results from Fitting Three Alternative Models to the Data

Factor	Model	Estimates h^2	c^2	i^2	b	e^2	h_b^2	χ^2	RMSEA	UL
E	*h,c,e*	.57	.00	.00	.00	.43	.57	125.38	.09	.11
	h,i,e	.04	.00	.57	.00	.39	.61	101.62	.08	.10
	h,c,e,b	.82	.00	.00	−.18	.23	.82	99.28	.08	.10
A	*h,c,e*	.51	.00	.00	.00	.49	.51	48.88	.04	.06
	h,i,e	.36	.00	.17	.00	.47	.53	47.68	.04	.06
	h,c,e,b	.58	.13	.00	−.12	.35	.58	47.21	.04	.06
C	*h,c,e*	.52	.00	.00	.00	.48	.52	38.21	.03	.05
	h,i,e	.40	.00	.13	.00	.47	.53	37.44	.03	.05
	h,c,e,b	.51	.22	.00	−.15	.32	.51	36.82	.03	.05
N	*h,c,e*	.58	.00	.00	.00	.42	.58	39.39	.03	.05
	h,i,e	.31	.00	.30	.00	.39	.61	34.59	.02	.05
	h,c,e,b	.74	.00	.00	−.10	.31	.74	34.72	.03	.05
O	*h,c,e*	.56	.00	.00	.00	.44	.56	32.00	.02	.04
	h,i,e	.36	.00	.22	.00	.41	.58	30.48	.02	.04
	h,c,e,b	.53	.38	.00	−.22	.22	.53	29.29	.01	.04

Note. E, A, C, N, O are the Big Five factors Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. The three models are, respectively, a simple model with genes (*h*), common environment (*c*), and unshared environment (*e*); a model with nonadditive genetic effects (*i*), and a model with contrast effects (*b*). χ^2 = goodness-of-fit chi square for the model in question. RMSEA = root mean square error of approximation. UL = upper limit of 90% confidence interval for RMSEA.

The *h,c,e,b* model estimates a twin interaction parameter *b*. It proves to be small and negative in the case of all five factors, consistent with a modest effect of contrast between the twins. The admission of this parameter still leaves the c^2 estimate at zero for Extraversion and Neuroticism, but allows for a positive estimate of the effect of shared family environment (13% to 38%) for Agreeableness, Conscientiousness, and Openness. The estimate of the broad heritability drops a bit for O, rises for E, A, and N, and stays about the same for C. Thus there is more difference in heritability across the five factors with the third model than with the first two: highest for Extraversion and Neuroticism (82% and 74%), intermediate for Agreeableness (58%), and lowest, though still substantial, for Conscientiousness and Openness (51% and 53%). Note that for this model the totals of genetic and environmental variances exceed 1.0. This is because negative covariance terms involving the path *b* combine with the simple totals in a complete variance accounting. But it also means that some caution should be exercised in interpreting heritabilities from this model—they are one step more hypothetical than in the other two.

DISCUSSION

There is a simple model of genetic and environmental influences that often fits behavior-genetic data on personality: genes plus individual environment, with little or no effect of shared environment. In the large National Merit twin sample, this model provides an acceptable fit to the common factor underlying the three measures of each of the Big Five traits. The estimated path values under this model imply that 51% to 58% of individual difference variation along the Big Five dimensions is genetic in origin, 42% to 49% is due to experience unique to the individual, to temporary situational factors, and to gene-environment interaction, and none is due to effects of environment shared by the twins.

We can complicate this view a little. When we look at the variance distinctive to each of the measures, as opposed to what they have in common, there seem to be shared effects of the twins' environments for the CPI measures of Agreeableness and Conscientiousness. The latter is consistent with results from a different approach using these same data, which found a significant shared environmental component for a ''norm-favoring'' scale of the CPI (Loehlin & Gough, 1990). These results are also in line with the finding of shared environment effects in four residual facet scales of the Revised NEO (Jang et al., in press). Two such facets lay in the Agreeableness domain (Altruism and Modesty), two in the Conscientiousness domain (Deliberation and Achievement Striving). Note that the CPI Big Five measures of the present study were derived via the NEO. Also consistent with the results of the Jang et al. study are significant genetic components to the measure-specific variance in several of the domains. In the present study, this might in part reflect differential involvement of the different facets in the different measures.

Another way we can complicate the simple view is by looking at models that incorporate nonadditive genetic effects or interactions between the twins. For A, C, and O neither of these alternative models produced a significant improvement in fit. For Extraversion, either one produced a significantly better fit, although still not an outstandingly good one. For Neuroticism, either alternative model led to a significant improvement of an already very good fit. The nonadditive genetic model assumed the presence of epistasis: interactions among multiple gene loci in their effects on a trait. The modeling suggested that perhaps half of the genetic effect on N and most of that on E could be of this type. The twin interaction model resulted in a small intertwin path that was negative for all five factors, in the range −.10 to −.22. The effect was one of contrast, an exaggeration of the difference between the twins.

The two alternative models tended to be about equally effective. They may be considered to represent alternative interpretations of essentially the same facts: DZ correlations less than half MZ correlations. This ambiguity could perhaps be most directly addressed by data from DZ twins reared apart. For them, correlations could be made lower by epistasis but not by the twins being contrasted with each other. Some data on DZ twins reared apart exist for the traits of Extraversion and Neuroticism. For Extraversion, the data tend to support a nonadditive genetic rather than a contrast model. DZ twins who had been reared apart from an early age still showed correlations considerably less than half those of MZ twins. This was true in three studies: one in Finland (Langinvainio, Kaprio, Koskenvuo, & Lönnqvist, 1984), one in Sweden (Pedersen, Plomin, McClearn, & Friberg, 1988), and one in Minnesota (Tellegen, Lykken, Bouchard, Wilcox, Segal, & Rich, 1988).

The situation for Neuroticism is less clear. In the Finnish study, DZs reared apart were found to be correlated less than half as much as MZs, and to be correlated much like DZs reared together. Both results are consistent with a hypothesis of nonadditive genetic effects. However, the other studies did not find DZ correlations for Neuroticism to be less than half MZ correlations to begin with, rendering moot an appeal to alternative models to explain such a difference. In the Swedish study, the correlation for DZ twins reared apart was slightly higher than that for DZ twins reared together (.28 vs .24), in line with a contrast effect, but the difference was in the opposite direction in the Minnesota study (.29 and .41). No general conclusion concerning alternative models seems justified for this trait.

The differences in h^2 and c^2 among the measure-specific components appear to be real, at least between the CPI measure and the others. The checklist and trait self-ratings scales were derived by workers in one of the two major Big Five traditions (Saucier & Goldberg, 1996), the so-called lexical tradition, which focuses on the Big Five as dimensions of person description in natural languages. The other tradition emphasizes the Big Five as fundamental dimensions of personality, and is mainly centered on work with the NEO-PI of Costa and McCrae, from which the CPI scales of the present study were derived. Thus we may have evidence of differences between these approaches at the measure-specific level, along with a consistent core.

The analysis by males and females separately suggested that sex differences in either general or specific heritabilities were not of substantial importance, but measurement differences might be for N and E.

In summary, if we take what various measures of the Big Five have in common, and fit simple behavior-genetic models, we obtain a simple result: genes accounting for something over half of individual differences along all five dimensions, with the rest presumably due to the effects of environmental inputs that are distinctive to each individual, temporary situational effects, and genotype-environment interactions (the last three in unknown proportions). The effect of environmental variables that act in the same way on both twins is estimated as zero for all five domains. We must emphasize that this does not mean that environment is unimportant for the development of personality, or as a source of individual differences. It merely means that whatever happens to individuals that makes a lasting difference is mostly independent of their families, or depends on their genes, or has effects that are unique to the individual.

For two of the Big Five, Extraversion and Neuroticism, more complex models involving gene-gene or twin-twin interactions fit somewhat better, although our data are not sufficient to distinguish between these alternatives. Moreover, the aspects of the Big Five measures beyond their common core are to some extent independently influenced by the genes, and in the case of the CPI scales for Agreeableness and Conscientiousness, perhaps by shared environments also.

What does this tell us about the optimum level at which to carry out behavior-genetic analyses of personality? This issue was considered by Loehlin (1992, chapter 4). He concluded that analyses below the level of the Big Five contribute significant added information. The finding in the present study of differences between common and measure-specific variance, and the finding of Jang et al. (1998) of differences among residualized Big Five facets, provide further evidence for this. Recent studies, including this one, make a case for saying that all five broad factors are substantially heritable and largely unaffected by shared environmental influences. However, the Big Five factors, or at least commonly-used measures of them, appear not to be monolithic in how they are influenced by the genes and the environment. Further research could profitably turn to a closer examination of specific subtraits, using multiple measures and observers.

Finally, it should not be forgotten that our results derive from a particular population—twin pairs from among the U.S. high school juniors who took the National Merit Scholarship Qualifying Test in the year 1962. If this study stood alone, the generality of its results might be in question. Since it does not, this specificity may be considered a virtue rather than a limitation. In particular, these data were gathered some decades earlier than those of other Big Five twin studies, and from a population homogeneous in age and education. Insofar as its results agree with the others, the overall generality of the findings increase. Further confirmation using nontwin behavior genetic methods, such as adoption or family studies, remains desirable.

APPENDIX A1

Covariance Matrices for Big Five Factors for MZ Twins (below Diagonal) and DZ Twins (Above Diagonal)

	TR1	CPI1	ACL1	TR2	CPI2	ACL2	
			Extraversion				
	.998	.6670	.7706	.0140	.1048	−.1279	TR1
		1.048	.5614	.1691	.3165	.0206	CPI1
TR1	.968		1.107	−.0773	.0317	−.0626	ACL1
CPI1	.6545	.956		1.130	.8250	.7634	TR2
ACL1	.6260	.4960	.877		1.072	.5529	CPI2
TR2	.4507	.4503	.3049	.964		1.082	ACL2
CPI2	.4452	.5822	.3274	.6819	.972		
ACL2	.3890	.3860	.3727	.6379	.5736	1.021	
			Agreeableness				
	.992	.2376	.4922	.0607	−.1035	.0537	TR1
		.742	.2803	.1020	.3065	.1013	CPI1
TR1	.910		1.025	.1681	.0845	.1872	ACL1
CPI1	.3761	1.038		1.148	.3907	.6120	TR2
ACL1	.5234	.3532	.993		1.114	.3683	CPI2
TR2	.2915	.1673	.2482	.930		.999	ACL2
CPI2	.2361	.4800	.2514	.3473	1.056		
ACL2	.2504	.1959	.2897	.4983	.3857	.985	
			Conscientiousness				
	1.104	.4403	.5426	.2230	.1487	.0729	TR1
		1.047	.5505	.1088	.3448	.0557	CPI1
TR1	.874		1.003	.1566	.1580	.1323	ACL1
CPI1	.3548	.976		1.100	.4778	.5807	TR2
ACL1	.5298	.4176	1.038		1.003	.3988	CPI2
TR2	.3860	.2189	.2423	.963		.911	ACL2
CPI2	.2665	.5249	.2749	.3672	.991		
ACL2	.2214	.2087	.3821	.5157	.4021	1.027	
			Neuroticism				
	.996	.4151	.6523	.1730	.0537	.0974	TR1
		.950	.5302	.1602	.2438	.1127	CPI1
TR1	1.004		1.063	.1127	.1098	.0560	ACL1
CPI1	.4715	.992		1.068	.4951	.6402	TR2
ACL1	.5694	.5029	.992		1.103	.5168	CPI2
TR2	.4152	.2331	.2720	.920		.966	ACL2
CPI2	.3493	.5280	.3792	.4655	1.018		
ACL2	.3613	.3204	.4431	.5540	.5955	1.034	
			Openness				
	1.076	.2214	.4477	.1971	.1113	.0881	TR1
		.953	.3157	.0537	.2772	.0199	CPI1
TR1	.946		1.092	.0803	.1329	.0788	ACL1
CPI1	.2364	.984		1.054	.2834	.4259	TR2
ACL1	.3514	.2526	1.039		1.068	.1867	CPI2
TR2	.3762	.0798	.1897	.972		.929	ACL2
CPI2	.1091	.4857	.2477	.1786	1.000		
ACL2	.1921	.1267	.3586	.3709	.2374	.944	

Note. TR = trait ratings, CPI = personality inventory items, ACL = adjective check list; 1,2 = twins.

REFERENCES

Bergeman, C. S., Chipuer, H. M., Plomin, R., Pedersen, N. L., McClearn, G. E., Nesselroade, J. R., Costa, P. T., Jr., & McCrae, R. R. (1993). Genetic and environmental effects on Openness to Experience, Agreeableness, and Conscientiousness: An adoption/twin study. *Journal of Personality,* **61,** 159–179.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin,* **117,** 187–215.

Costa, P. T., Jr., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology,* **55,** 258–265.

Costa, P. T., Jr., & McCrea, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology,* **41,** 417–440.

Eaves, L. J. (1988). Dominance alone is not enough. *Behavior Genetics,* **18,** 27–33.

Gough, H. G. (1957). *CPI manual*. Palo Alto, CA.: Consulting Psychologists Press.

Gough, H. G., & Heilbrun, A. B., Jr. (1965). *The Adjective Check List manual.* Palo Alto, CA: Consulting Psychologists Press.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd. ed.). Port Huron, MI: Research Psychologists Press.

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology,* **74,** 1556–1565.

John, O. P. (1990). The ''Big Five'' factor taxonomy: Dimensions of personality in the natural languages and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

John, O. P., & Roberts, B. W. (1993). *Measuring the Five-Factor Model on the Adjective Check List.* Technical Report, Institute of Personality and Social Research, University of California, Berkeley, CA.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Mahwah, NJ: Erlbaum.

Langinvainio, H., Kaprio, J., Koskenvuo, M., & Lönnqvist, J. (1984). Finnish twins reared apart. III: Personality factors. *Acta Genetica Medicae et Gemellologiae,* **33,** 259–264.

Loehlin, J. C. (1992). *Genes and environment in personality development.* Newbury Park, CA: Sage.

Loehlin, J. C., & Gough, H. G. (1990). Genetic and environmental variation on the California Psychological Inventory vector scales. *Journal of Personality Assessment,* **54,** 463–468.

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

Lykken, D. T. (1982). Research with twins: The concept of emergenesis. *Psychophysiology,* **19,** 361–373.

McCrae, R. R. (1989). Why I advocate the five-factor model: Joint factor analyses of the NEO-PI with other instruments. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 237–245). New York: Springer-Verlag.

McCrae, R. R., & Costa, P. T., Jr. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality,* **57,** 17–40.

McCrae, R. R., Costa, P. T., Jr., & Piedmont, R. L. (1993). Folk concepts, natural language, and psychological constructs: The California Psychological Inventory and the Five-Factor Model. *Journal of Personality,* **61,** 1–26.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto: Consulting Psychologists Press.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families.* Dordrecht, The Netherlands: Kluwer.

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

Price, R. A., & Vandenberg, S. G. (1980). Spouse similarity in American and Swedish couples. *Behavior Genetics,* **10,** 59–71.

Riemann, R., Angleitner, A, & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality,* **65,** 449–475.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the Five-Factor model. In Wiggins, J. S. (Ed.), *The Five-Factor model of personality: Theoretical perspectives* (pp. 21–50). New York: Guilford.

Shock, N. W., Greulich, R. C., Andres, R., Arenberg, D., Costa, P. T., Jr., Lakatta, E. G., & Tobin, J. D. (1984). *Normal human aging: The Baltimore Longitudinal Study of Aging* (NIH Publication No. 84–2450). Bethesda, MD: National Institutes of Health.

Steiger, J. H. (1989). *EZPath: A supplementary module for SYSTAT and SYGRAPH.* Evanston, IL: SYSTAT, Inc.

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

Waller, N. G. (in press). Evaluating the structure of personality. In C. R. Cloninger (Ed.), *Personality and psychopathology*. Washington, D.C.: American Psychiatric Press.