



# Genetic and environmental structure of adjectives describing the domains of the Big Five Model of personality: A nationwide US twin study

Wendy Johnson\* and Robert F. Krueger

*Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, USA*

Available online 15 January 2004

---

## Abstract

We applied multivariate models specifying genetic and environmental influences on adjectives describing each of the five personality domains specified in the Big Five Model of personality (BFM; Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness). We selected the specific models to partition the observed covariance among the adjectives describing each domain into genetic and environmental components in order to assess the etiologic basis for each domain's phenotypic coherence. The sample on which our analyses were based was part of the National Survey of Midlife Development in the United States (MIDUS). It consisted of 315 monozygotic and 275 same-sex dizygotic twin pairs. Results revealed both common and specific genetic and environmental influences for each domain, suggesting that all of the domains are etiologically complex. Models specifying the domains as latent phenotypic constructs fit more poorly than models suggesting more complex structures for all domains except Extraversion and Neuroticism. These results raise questions about the BFM as a coherent model of genetic and environmental influences on personality or, alternatively, about the etiological unity of latent phenotypic personality trait constructs beyond Extraversion and Neuroticism.

© 2003 Elsevier Inc. All rights reserved.

*Keywords:* Genetic and environmental influences; Personality structure; Big Five Model; Twin study

---

---

\* Corresponding author. Fax: 1-952-473-1998.

E-mail address: [john4350@tc.umn.edu](mailto:john4350@tc.umn.edu) (W. Johnson).

## 1. Introduction

Factor analysis has been used to develop many models of the structure of personality (e.g., Eysenck & Eysenck, 1985; McCrae & Costa, 1990; Tellegen, 1982). The one that has received the most attention in recent years, however, is the Big Five Model (BFM), originally identified in lexical studies (Goldberg, 1990; Saucier & Goldberg, 1996a). This model allocates personality variance among the domains Extraversion (E), Neuroticism (N), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). According to the premises of the lexical approach used to identify the model (Saucier & Goldberg, 1996b), these constructs refer to observable, surface characteristics (phenotypes) rather than to underlying causal properties such as genotypes.

In developing the NEO-PI-R to operationalize the BFM, Costa and McCrae (1995) have broadened the conceptualization of the five personality factors considerably beyond the lexical approach. They maintain that each of these domains represents a coherent personality construct. In describing the role the domains play in manifested behaviors, McCrae and Costa (1999) specifically assign them to the level of basic tendencies, thus distinguishing them from patterns of behavior, and from the plans, skills, and desires that lead to patterns of behavior: “They [the domains] are directly accessible neither to public observation nor to private introspection. Instead, they are deeper psychological entities that can only be *inferred* [italics theirs] from behavior and experience (p. 143).” In short, Costa and McCrae define them to be unified latent personality traits.

This definition has been accepted in the sense that there have been a number of analyses of the genetic and environmental influences on the Big Five. These analyses have generally estimated such influences scale by scale (e.g., Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; Loehlin, 1992; Loehlin, McCrae, Costa, & John, 1998; Riemann, Angleitner, & Strelau, 1997; Waller, 1999), and have shown highly consistent estimates of substantial genetic influences and little or no shared environmental influences on all five domains, providing evidence for a causal mechanism at least for the observable traits. Acceptance of the stronger definition of the domains as latent traits has not been universal, however. For example, Paunonen and Jackson (1996) suggested that each of the five domains actually consists of separate subdomains. Though they may be correlated, the argument is that these subdomains may not be closely enough related to be considered a unified personality trait.

Hofstee, de Raad, and Goldberg (1992) have elaborated considerably on this conception in the process of attempting to integrate the Big Five with the circumplex model, in which traits are characterized by their angular positions in a two-dimensional factor space. Their five-dimensional circumplex model provided clear evidence that the majority of trait terms load not onto single Big Five domains, but rather cross-load onto two Big Five domains. This would suggest that individual trait terms should be considered blends of two factors (as suggested by Saucier & Ostendorf, 1999). Hofstee, de Raad, & Goldberg thus concluded that there are nine potential bipolar facets for each of the Big Five: core traits representing that domain alone, without cross-loading, and additional facets representing that domain at a primary

level and another Big Five domain at a secondary level. This would suggest that it is unreasonable to expect the domains to represent unitary personality constructs based on primary loadings, as even the categorization based on primary loadings would consist of trait terms with substantial secondary loadings on other domains that are considered orthogonal.

Costa and McCrae (1998) used six methodological approaches to show that their Conscientiousness domain is unitary in nature. These six approaches were: (1) item content analysis, (2) definitions of psychological opposites, (3) examination of empirical correlates, (4) interpreting secondary and tertiary factor loadings, (5) identification of equivalents in specialized languages, and (6) case studies. These approaches, though thorough in that they help to clarify the observable, phenotypic structure of personality, cannot increase our understanding of its etiological underpinnings. If the Big Five domains are in fact latent personality traits, it is reasonable to expect that the covariance structures of their subdomains should each be caused by a set of genetic and/or environmental influences that act in a unified way on the latent trait. We assume that any etiological structure of any personality trait is going to be polygenic and polyenvironmental. But the individual genes may operate in concert to varying degrees, and the same is true of the environmental factors. Most studies that investigate genetic and environmental influences on a single trait rely on the assumption that that trait is influenced by genes and environmental factors operating in a unified manner to influence the trait. In order to test this assumption directly, however, it is necessary to make use of multivariate genetically informed studies. Such studies can thus be considered to be robust ways of comparing the conceptions of the Big Five as coherent latent personality traits and as loose organizational categories.

This kind of multivariate genetically informed approach was used by Heath, Eaves, and Martin (1989) to evaluate the unity of the structure of the items making up the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1985) scales. The EPQ consists of three substantive scales of 21–25 items each, measuring Extraversion, Neuroticism, and Psychoticism, as well as a validity or Lie scale. Heath et al. (1989) found that single common genetic and environmental factors could be extracted for Extraversion and Neuroticism, but that the Psychoticism items required two distinct genetic factors, which they labeled “paranoid attitudes” and “hostile behavior.” They interpreted their findings as providing evidence that Extraversion and Neuroticism are coherent concepts at an etiological level, but that Psychoticism is not.

Some multivariate analyses examining the coherence of the latent domains of the BFM have also been conducted. For example, Loehlin et al. (1998) found substantial genetic influences and little shared environmental influence on the variance common to three different Big Five personality indexes, a trait self-rating scale classified to fit the Big Five, a standard personality inventory fit to the Big Five using analysis of correlation with the NEO-PI-R, and an adjective check list. McCrae, Jang, Livesley, Riemann, and Angleitner (2001) factor analyzed phenotypic, genetic/familial, and residual covariance matrices of the NEO-PI-R, supplemented with cross-observer correlations. They found the hypothesized five-factor structure

in both the phenotypic and genetic/familial covariances, but not in the residual matrices. When the residual matrices were decomposed into nonshared environmental covariance and method covariance, only the method variance showed the five-factor structure. This suggested that the five-factor structure can be extracted from data at multiple levels of underlying structure, but that it cannot be extracted universally, and in particular it could not be extracted from data with little systematic underlying structure.

The multivariate approach was also used by Jang, Livesley, Angleitner, Riemann, and Vernon (2002) to conduct a multivariate analysis of the genetic and environmental influences on the five NEO-PI-R domains. Using both a sample from Canada and a sample from Germany, they fit a series of path analytic models to the data. The models were of two basic types. The first type was a model known as *independent pathways*, which specifies direct links to the subdomains from one or more additive genetic and shared and nonshared environmental influences common to all the variables, but no common underlying latent phenotype. This model could be expected to fit best if Hofstee et al.'s (1992) conception of the BFM as a series of linked circumplexes reflecting primary and secondary BFM loadings is an accurate description of the situation, as the genetic and environmental influences on the primary domain for each facet should be common to all facets in the primary domain, but the genetic and environmental influences on the domains associated with the secondary loadings should be different. The second type was a model known as *common pathways*. This model provides a more stringent test of Costa and McCrae's (1995) hypothesis of latent personality constructs in that it specifies that all of the covariation in a set of variables be mediated by a single latent phenotypic variable influenced by one set of additive genetic and shared and nonshared environmental influences. Additive genetic influences reflect the extent to which genotypes are transmitted directly from parents to offspring. Shared environmental influences (e.g., neighborhoods) affect all children within a family to the same degree and differentiate between families. Nonshared environmental influences (e.g., differential parental treatment) have different effects on individual family members. Error variance is also included with nonshared environmental influence because the nonshared environmental influence terms are the residuals after the effects of additive genetic and shared environmental influences have been estimated.

Jang et al. (2002) concluded that each BFM domain was made up of more than one common genetic factor and more than one common environmental factor, in keeping with Hofstee et al.'s (1992) conception of the BFM. That is, the domains did not appear to be unidimensional either on the genetic or the environmental level. At the same time, there was one common genetic factor that influenced all the facets in each of the E, N, and C domains. In addition, most of the facets also loaded on the other common factor(s). There were also many large genetic and environmental influences specific to individual facets, and the common pathways model failed to fit adequately. Jang et al. (2002) suggested that their results called into question the meaning of the concept of a higher-order phenotypic trait. They pointed out that the Big Five higher-order personality traits may not exist as unitary psychological entities, but rather, may simply be useful heuristic devices that,

at best, can be associated with independent activity of many genes and multiple environmental factors.

The purpose of the current study was to extend the existing sparse literature on the genetic and environmental structure underlying the BFM by examining it in adjective descriptors specifically designed to assess the Big Five in a nationwide sample of twin pairs in the United States. The sample was recruited by screening a nationally representative sample of 50,000 households for twins. To our knowledge, it was the first twin sample to be recruited in this manner. Participants in the full twin sample ranged in age from 25 to 74 years, and included residents of every state in the continental United States and the District of Columbia, roughly distributed in proportion to population by state. We conducted our study by applying a series of multivariate path analytic models, including the common and independent pathways models, to the BFM data.

## 2. Method

### 2.1. Sample

The twin sample used in this study was gathered as part of the MacArthur Foundation Survey of Midlife Development in the United States (MIDUS). Two research organizations, ICR/AUS Consultants and Bruskin Associates, recruited twin pairs by making telephone calls and asking respondents whether they or any of their immediate family members were members of intact twin pairs. About 50,000 households, constituting a representative national sample, were screened in this manner. The 14.8% of respondents who reported the presence of a twin in the family were then asked whether it would be acceptable for the research team to contact the twins to solicit their participation in the survey. The 60% of the respondents who gave such permission were referred to the MIDUS recruitment process (see Kendler, Thornton, Gilman, & Kessler, 2000 and Kessler, Gilman, Thornton, & Kendler, in press, for additional details).

Both members of each twin pair participating in the study met both overall study eligibility criteria and criteria specific to the twin sample. These criteria included: (1) being at least first degree relatives of the original contact or his or her spouse or partner, (2) being between the ages of 25 and 74, (3) having a residential telephone number, (4) living in the continental United States, (5) speaking English, and (6) being mentally and physically able to complete the interview and questionnaires. The base MIDUS twin sample resulting from this process consists of 998 pairs.

Zygosity for each twin pair was determined using self-report questions regarding information such as similarity of eye and hair color and degree to which others were confused as to their identity during childhood. Such techniques are generally more than 90% accurate (Lykken, Bouchard, McGue, & Tellegen, 1990), though some pairs in this sample were not considered classifiable on the basis of the self-reports. We made use of the 590 same-sex pairs for which we had both personality inventory data and zygosity data for the current study, resulting in 141 monozygotic (MZ)

male pairs, 174 MZ female pairs, 107 dizygotic (DZ) male pairs, and 168 DZ female pairs. We thus excluded 262 opposite-sex pairs, 16 pairs with missing or indeterminate zygosity information, and 130 pairs with incomplete or missing personality data from the full MIDUS twin sample of 998 pairs. Table 1 also shows demographic information for the sample we used for this study.

## 2.2. Personality inventory

The self-administered mailed questionnaire booklets that were part of the MIDUS survey included a personality measure based on the BFM (Lachman & Weaver, 1997). The measure consisted of 25 adjectives chosen from existing trait lists and inventories (e.g., Goldberg, 1992; John, 1990; Trapnell & Wiggins, 1990) to reproduce the BFM. Based on the responses of participants in a pilot study ( $N = 1000$ , age range 30–70 years), the adjectives that appeared most consistently as trait markers and had the highest item-to-total correlations or factor loadings were identified. The resulting instrument had four to seven adjectives intended to index each Big Five domain, rated on 4-point scales. It had satisfactory scale consistencies for most of the domains ( $\alpha = .74, .78, .80, .58$ , and  $.77$  for Neuroticism, Extraversion, Agreeableness, Conscientiousness, and Openness, respectively), though the  $\alpha$  for Conscientiousness was somewhat low. John (1990) has suggested that different measurement formats of the BFM (e.g., adjectives vs. full statements) are basically

Table 1  
Sample demographic information

Category	Percent
Sex	
Male	42.4
Female	57.6
Race	
White	91.8
Black/African American	4.1
Others	1.9
Not reported	2.2
Age	
25–34	21.7
35–44	28.8
45–54	26.4
55–64	14.4
65–74	8.5
Education	
Under 12 years	9.5
12 years	32.0
Over 12 years	58.1
Not reported	0.3
Marital status	
Married	74.7
Not married	25.3

Table 2  
Means and standard deviations of Big Five Model adjectives

Factor	Mean	SD
<i>Extraversion</i>		
Outgoing	1.94	.84
Lively	1.82	.74
Talkative	2.06	.90
<i>Neuroticism</i>		
Moody	2.81	.85
Worrying	2.46	.92
Nervous	2.78	.94
Not calm	3.00	.78
<i>Agreeableness</i>		
Helpful	1.40	.57
Friendly	1.36	.55
Warm	1.51	.63
Caring	1.34	.54
Softhearted	1.59	.71
Sympathetic	1.52	.65
<i>Conscientiousness</i>		
Organized	1.91	.83
Responsible	1.30	.53
Hardworking	1.28	.53
Not careless	1.72	.71
<i>Openness</i>		
Creative	2.14	.87
Imaginative	1.98	.81
Intelligent	1.74	.65
Curious	1.73	.74
Active	1.70	.72
Broad-minded	1.94	.79
Sophisticated	2.59	.87
Adventurous	2.15	.84

*Note.* Adjectives were rated on a scale of 1–4 indicating “how well each of the following describes you: a lot, some, a little, not at all,” with 1 indicating “a lot.”

equivalent. The means and standard deviations of the BFM adjective scores are shown in Table 2.

### 2.3. Analytical approach

#### 2.3.1. Factor analysis of MIDUS adjectives

We began our analysis by conducting a factor analysis of the adjectives in the MIDUS twin data. Because we were planning to conduct a multivariate analysis of genetic and environmental influences on adjectives within each Big Five domain, we wanted to be sure that we assigned the adjectives to the domains with which they had the most common variance in the data set.

### 2.3.2. Age adjustment

Because co-twins are the same age, age effects act to increase twin similarity (McGue & Bouchard, 1984). To correct for this in the analysis of genetic and environmental influences on the twin data, we regressed out the effects of age and age<sup>2</sup> on each factor descriptor scale prior to fitting our biometric models. This also acted to normalize the variables.

### 2.3.3. Biometric modeling

The standard univariate quantitative genetic model is based on the understanding that the observed phenotypic variance ( $V_p$ ) is a linear additive function of genetic ( $V_g$ ) and shared ( $V_s$ ) and non-shared ( $V_n$ ) environmental variance, respectively. Symbolically,

$$V_p = V_g + V_s + V_n.$$

Under this model, the non-shared environmental variance represents residual variance not explained by either of the other two sources. The non-shared environmental component also includes variance attributable to measurement error. Genetic variance can be additive ( $V_{ga}$ ) in the sense that if multiple genes influence the trait, they do so independently of each other. It can also be nonadditive ( $V_{gn}$ ), reflecting dominance and other polygenic effects. The expected covariance between any two members of a twin pair as a function of the variance components given above can thus be specified as,

$$\text{COV}_{(MZ)} = V_{ga} + V_{gn} + V_s,$$

$$\text{COV}_{(DZ)} = .5 * V_g + .25 * V_{gn} + V_s.$$

(Nonadditive genetic variance is usually modeled as the effects on variance of dominant genes.) Models containing terms to reflect both forms of genetic variance and both forms of environmental variance are not identified using data from twins reared together; thus models including only three of the four forms of variance at a time must be fit in stages. We note that there is little evidence for shared environmental effects on personality (Bouchard & Loehlin, 2001; Finkel & McGue, 1997; Tellegen et al., 1988). We fit models both including and excluding parameters for shared environmental effects. In all cases the estimated shared environmental effects were negligible, and the models excluding them fit better. We note that there is evidence for some nonadditive effects on personality, though the particular effects do not generally replicate (see Bouchard & Loehlin, 2001; Bouchard & McGue, 2003 for recent reviews). We also fit models including nonadditive genetic effects. Though there were a few cases for which the nonadditive genetic effects were significant for models of specific types for specific factors and some cases for which the model of a particular type for a particular factor including the nonadditive genetic effects fit significantly better (as measured by the  $\chi^2$  statistic) than the analogous model without this parameter for that factor, this was never consistently the case across all three models fit for a factor and all fit statistics. In addition, there was never a case in which the inclusion of this parameter made any difference in the determination of



which of the three models was preferred for that factor. We thus report only the results for the models including only additive genetic and nonshared environmental parameters.<sup>1</sup>

The standard univariate model can be extended to multivariate situations by modeling the covariance between one twin's score on one variable and the other twin's score on another variable in a manner directly analogous to the univariate case. There are three general classes of path analytic models that can be fit to multivariate twin data.

The first is the *Cholesky model*, which can be used to establish baseline parameter estimates of genetic and environmental influences. We can think of the BFM trait score consisting of the sum of the ratings on the adjectives loading on that domain (e.g., extraversion) as a phenotype and the rating on each adjective as one aspect of this phenotype. Thus, being “outgoing,” “lively,” and/or “talkative” can be considered to be distinct aspects of extraversion and individuals may differ in their relative proportions of each as well as in their overall levels of the phenotype. Implementation of this model relies on the fact that any positive definite covariance matrix can be decomposed (uniquely, except for transformations of sign) into the product of a lower triangular matrix and its transpose. This can be done separately for the genetic and environmental portions of the phenotypic variance described above. The model is represented graphically as shown in Fig. 1. It imposes no underlying structure on the genetic and environmental influences, and simply recounts the extent of their interrelationships. The first latent factor of each type (genetic or environmental) loads on all the adjectives, the second on all adjectives except the first, the third on all adjectives except the first two, and so on. The order of the adjectives is arbitrary, and the measurements of interrelationships would be the same no matter what order was used. Because the model freely estimates genetic and environmental influences, imposing no latent genetic or environmental structure on the data, when it fits best it implies that the phenotype should be broken down into subtraits at the level of genetic and environmental influences. Thus, for example, if the Cholesky model fit best

---

<sup>1</sup> For Openness, though there was no significant difference in model fit between the models including and excluding nonadditive genetic variance for the Cholesky and Common Pathways models, the Independent Pathways model including nonadditive genetic variance fit significantly better than the analogous model excluding such variance ( $\chi^2 = 338.84$ , 224 *df* vs.  $\chi^2 = 452.19$ , 256 *df*,  $p < .001$ ). Both Independent Pathways models fit better than their Common Pathways counterparts, however ( $\chi^2 = 338.84$ , 224 *df* vs.  $\chi^2 = 590.11$ , 245 *df*,  $p < .001$  for the models including nonadditive genetic variance and  $\chi^2 = 452.19$ , 256 *df* vs.  $\chi^2 = 607.75$ , 262 *df*,  $p < .001$  for the models excluding it), and BIC for the Independent Pathways model including nonadditive genetic variance was  $-1090.35$ , so the Independent Pathways model excluding such variance was preferred. A similar situation existed for Agreeableness. The Independent Pathways model including nonadditive genetic variance fit significantly better than the analogous model excluding such variance ( $\chi^2 = 324.07$ , 132 *df* vs.  $\chi^2 = 384.31$ , 144 *df*,  $p < .001$ ), but both Cholesky models fit significantly better than their corresponding Independent Pathways models ( $\chi^2 = 184.93$ , 105 *df* vs.  $\chi^2 = 324.07$ , 132 *df*,  $p < .001$  for the models including nonadditive genetic variance and  $\chi^2 = 187.90$ , 126 *df* vs.  $\chi^2 = 384.31$ , 144 *df*,  $p < .001$  for those excluding it), and the respective BIC's were  $-518.18$  and  $-534.42$ , so the Cholesky model excluding nonadditive genetic variance was preferred.

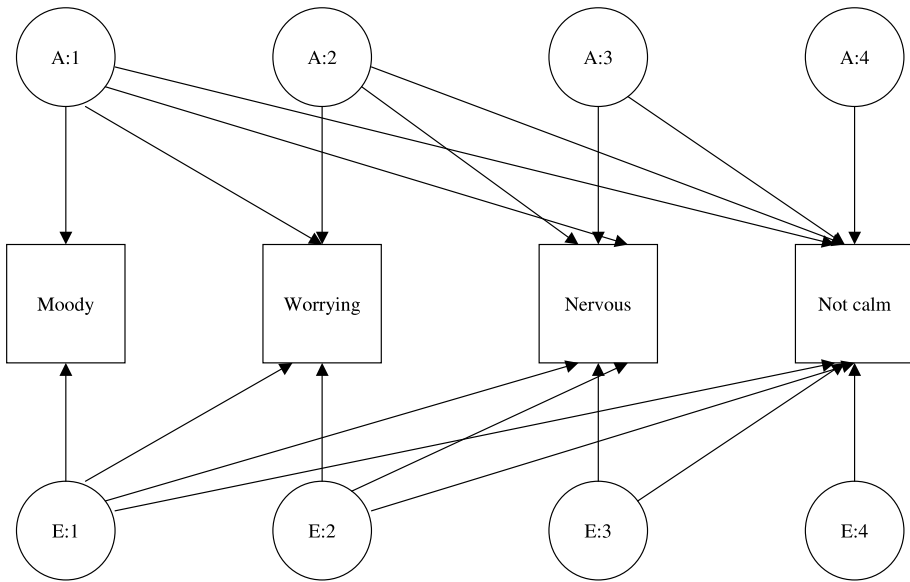


Fig. 1. Multivariate Cholesky model relating genetic and environmental influences across factor descriptors. A refers to genetic influence and E to non-shared environmental influence. Neuroticism is used as an example.

for extraversion, we could conclude that the etiology of being “outgoing” is separable at some genetic and/or environmental level from that of being “talkative.”

The second class of model is the independent pathways, shown in Fig. 2. Instead of a series of latent genetic and environmental factors as in the Cholesky, this model specifies direct links to the adjectives from a single set of additive genetic and environmental influences common to all the adjectives, as well as genetic and environmental influences specific to each adjective. For extraversion, this model would imply that common genetic and environmental influences would affect each of “outgoing,” “lively,” and “talkative,” but that these three indicators of extraversion are independent at a more fine-grained level. It imposes an intermediate level of structure on the data, as it is based on the assumption of a single set of genetic and environmental influences on all the adjectives, but does not specify that the adjectives form a single coherent latent phenotype. We expected this model to fit best if Hofstee et al.’s (1992) conception of the BFM as descriptive categories of traits is correct.

The third class of models is the common pathways. In this case, this model is based on the assumption that the covariation among adjective scores is determined by a single phenotypic latent variable, and we expected the model to fit best if Costa and McCrae’s, 1995 conception of the BFM domains as latent personality traits is correct. The latent variable has additive genetic and environmental influences, with direct phenotypic paths to each adjective score. As with the independent pathways model, there are also genetic and environmental influences specific to each adjective.

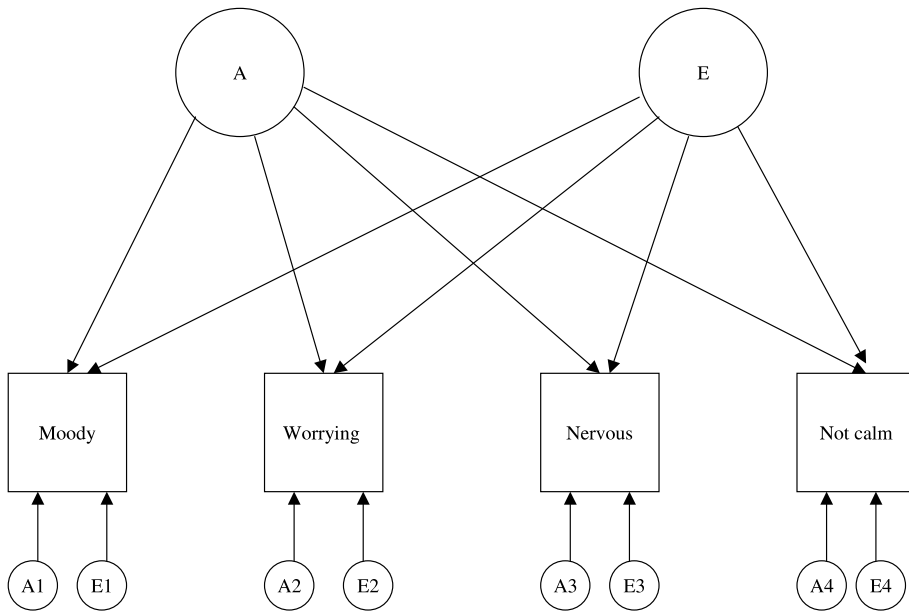


Fig. 2. Multivariate independent pathway model relating genetic and environmental influences across factor descriptors. A refers to genetic influence and E to non-shared environmental influence. Neuroticism is used as an example.

The basic model is shown in Fig. 3. For neuroticism, this model would imply that the four measures reflect different aspects of a single coherent underlying phenotype. Throughout, we conducted our analyses of twin covariance matrices using maximum likelihood estimation as operationalized in the computer program Mx (Neale, 1997). We analyzed males and females separately, comparing models with parameters separately estimated for each sex to those constraining parameters equal across sex. As the models with parameters estimated separately for each sex did not fit significantly better in any case, we only report those with parameters constrained equal across sex. We double-entered the twin data and adjusted the degrees of freedom accordingly to remove the effects of ad hoc variance differences between Twin 1 and Twin 2. We note that any assignment of order to the twin pairs is arbitrary and can result in such ad hoc variance differences. Double-entry removes these ad hoc variance differences effectively without changing the underlying structural relationships.

We assessed model fit using the  $\chi^2$  statistic, Akaike's Information Criterion (AIC; Akaike, 1983), and the Bayesian Information Criterion (BIC; Raftery, 1995). The  $\chi^2$  statistic does not reflect model parsimony, and tends to produce significant (ill-fitting) results when sample sizes are large (Raftery, 1995). In addition, it is only possible to use this statistic to compare the fit of nested models, and the Cholesky and independent and common pathways models do not form such a nested set. For these two reasons we calculated fit statistics AIC and BIC as well. AIC is defined as the  $\chi^2$  statistic for the model, minus two times the degrees of freedom. Smaller or more neg-

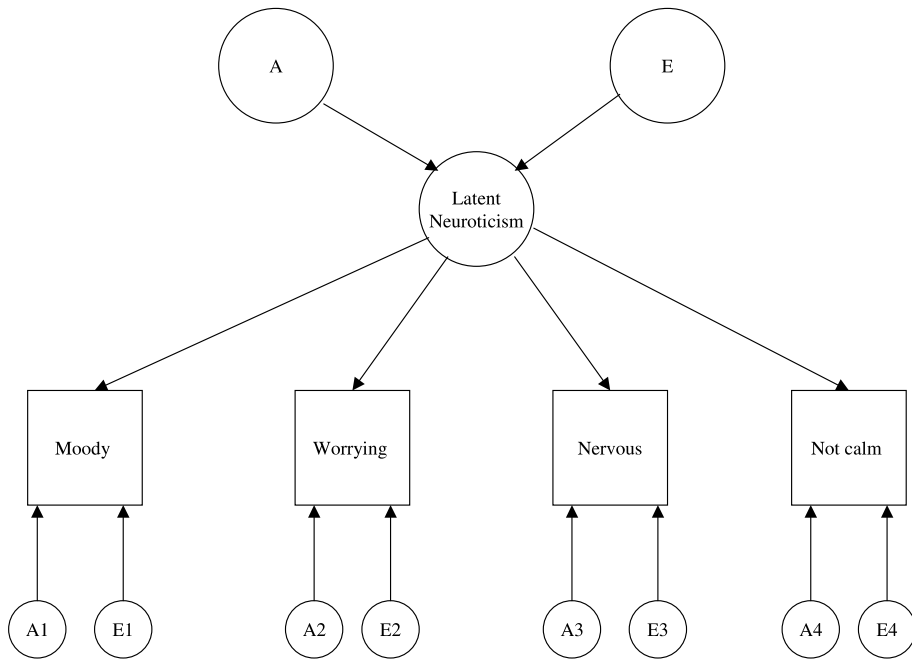


Fig. 3. Common pathway model showing genetic and environmental influences across factor descriptors. A refers to genetic influence and E to non-shared environmental influence. Neuroticism is used as an example.

ative AIC's are preferred, reflecting the balance between model parsimony and reproduction of observed data. BIC is defined as the  $\chi^2$  statistic for the model minus the product of the degrees of freedom and the natural log of the sample size. We used the overall number of pairs as the sample size for BIC, as the multivariate twin model relies on the covariance between twins' scores on different variables, making the pair the unit of analysis for structural modeling of twin data. Smaller or more negative BIC's are also preferred, reflecting model parsimony to a greater degree than AIC, particularly for larger sample sizes. Both AIC and BIC can be used to compare models whether nested or not. We tended to rely most heavily on BIC in the evaluation of our results as it places a stronger emphasis on model parsimony. Of the three models tested, the common pathway model is the most parsimonious; hence, all else being equal, BIC will prefer the common pathway model—the model that corresponds to the greatest degree of underlying structure.

The use of the linear additive variance decomposition formulas on which both univariate and multivariate genetic models are based relies on several assumptions. The first assumption is that twins (both MZ and DZ) are representative of the population as a whole for the trait in question. For personality traits, this appears to be the case (Johnson, Krueger, Bouchard, & McGue, 2002). Second, we assume that MZ twins share imposed trait-relevant environmental influences to the same degree as DZ twins. For example, violations of this assumption would take place if MZ

twins were more likely than DZ twins to be placed in the same school classroom, to be dressed alike, or to be placed by their parents in the same sports programs and these imposed environments resulted in greater MZ relative to DZ similarity in personality in adulthood. Violations of this assumption would not take place if MZ twins were to be more likely to choose similar friends, or to choose the same school elective classes or sports programs themselves because such choices could be manifestations of genetic factors (Krueger, Markon, & Bouchard, 2003). Numerous attempts have been made to uncover circumstances in which this assumption does not hold, with generally negative results (e.g., Borkenau, Riemann, Angleitner, & Spinath, 2002; Loehlin & Nichols, 1976).

Another assumption is that there is no assortative mating (meaning that the parents of the twins were not similar) for the traits in question. This assumption is reasonable for personality measures because there is only very modest marital resemblance (with coefficients ranging from .10 to .20) for most personality-related traits (Price & Vandenberg, 1980; Tellegen et al., 1988), and such relatively small values are unlikely to affect estimates of genetic and environmental influence substantially.

### 3. Results

#### 3.1. General results

##### 3.1.1. Factor analysis of the MIDUS adjectives

To evaluate the phenotypic structure of our personality measure, we made use of ordinary least squares factor extraction with varimax rotation specifying five factors, as operationalized in the computer program CEFA (Browne, Cudeck, Tateneni, & Mels, 2001). We used varimax rotation because it is the method most commonly applied in other studies of the BFM.

The major loadings and communalities in the factor solution are shown in Table 3. The communalities were reasonable, with almost 70% of them in the .4–.7 range. The five-factor structure was clearly visible and could be readily interpreted as the BFM, though the three items loading on Extraversion certainly could be viewed as a very narrow version of the trait, and the Agreeableness factor seemed to include some of the elements missing from Extraversion. Using .3 as the criterion for a substantial loading, all the adjectives except “not careless” had substantial loadings on at least one factor, and only four adjectives had such loadings on more than one factor. In carrying out our biometric modeling, we assigned each adjective to the factor on which it had the highest loading. This had the result that “friendly” and “warm” were assigned to Agreeableness, instead of to Extraversion as had been intended.

##### 3.1.2. Twin correlations

The MZ and DZ intraclass twin correlations are shown in Table 4. As would be expected given results from other studies of the BFM, all but one (“sympathetic”) of

Table 3  
Factor analysis solution of MIDUS personality adjectives under the Big Five Model

Factor	Extraversion	Neuroticism	Agree- ableness	Conscien- tiousness	Openness	Community
Outgoing	.74					.64
Lively	.60				.31	.55
Talkative	.59					.45
Moody		.59				.36
Worrying		.77				.62
Nervous		.80				.66
Not calm		.47				.38
Helpful			.50			.43
Friendly	.48		.54			.57
Warm	.41		.68			.67
Caring			.73			.61
Softhearted			.62			.41
Sympathetic			.66			.47
Organized				.52		.31
Responsible				.64		.48
Hardworking				.46		.32
Not careless				.28		.15
Creative					.62	.41
Imaginative					.72	.54
Intelligent					.47	.32
Curious					.58	.38
Active				.33	.44	.42
Broad- minded					.39	.22
Sophisticated					.42	.47
Adventurous					.57	.40

*Note.* Extraction method was unweighted least squares with varimax rotation, five factors extracted. All loadings above .30 are shown, as well as the highest loading for “not careless,” which was .28.

the MZ correlations exceeded the DZ correlations, suggesting genetic influences on the adjective scores. The MZ correlations ranged from .15 to .42. The low MZ correlations were concentrated in Agreeableness and Conscientiousness, with a few in Openness as well. The DZ correlations ranged from .02 to .28.

### 3.1.3. Biometric modeling

The model-fitting statistics for the three multivariate models applied to each of the domains of the BFM are shown in Table 5. There were some patterns in the results common to the models for all five BFM domains, and some differences among them. We outline the common patterns first. Relying on BIC, which emphasizes model parsimony to a very high degree and therefore provides a more liberal test of the latent trait hypothesis, the common pathways model fit best for Extraversion and Neuroticism, the independent pathways model fit best for

Table 4  
 Intraclass twin correlations of the Big Five Model adjectives

Factor	MZ ( <i>N</i> = 315 pairs)	DZ ( <i>N</i> = 275 pairs)
<i>Extraversion</i>		
Outgoing	.38	.12
Lively	.35	.09
Talkative	.42	.15
<i>Neuroticism</i>		
Moody	.35	.16
Worrying	.38	.11
Nervous	.42	.12
Not calm	.40	.28
<i>Agreeableness</i>		
Helpful	.26	.14
Friendly	.29	.14
Warm	.27	.06
Caring	.20	.02
Softhearted	.15	.08
Sympathetic	.21	.27
<i>Conscientiousness</i>		
Organized	.37	.13
Responsible	.24	.08
Hardworking	.24	.16
Not careless	.22	.17
<i>Openness</i>		
Creative	.34	.17
Imaginative	.30	.11
Intelligent	.42	.07
Curious	.22	.17
Active	.25	.13
Broad-minded	.18	.17
Sophisticated	.31	.22
Adventurous	.35	.13

*Note.* MZ is monozygotic; DZ is dizygotic. Standard errors for these correlations ranged from .05 to .06.

Conscientiousness and Openness, and the Cholesky model fit best for Agreeableness. (Note that AIC always indicated that the Cholesky model provided the best fit.) The common pathways model imposes the structure of a latent phenotypic personality construct on the data and should provide the best fit if the latent trait hypothesis is accurate; the independent pathways model imposes an intermediate level of latent structure and should provide the best fit if the loose organizational structure hypothesis is accurate; and the Cholesky imposes no specific latent structure. We show the parameter estimates and variance components from the common pathways model for Extraversion and Neuroticism in Table 6, and the analogous results from the independent pathways model for the other three BFM domains in Table 7. We show the independent pathways parameters for

Table 5  
Model-fitting statistics for the adjectives of the domains of the Big Five Model

Factor and model	$\chi^2$	<i>df</i>	<i>p</i>	AIC	BIC
<i>Extraversion</i>					
Cholesky	35.15	36	.51	–36.84	–194.53
Independent pathways	35.23	36	.51	–36.77	–194.45
Common pathways	37.22	37	.46	–36.78	–198.52
<i>Neuroticism</i>					
Cholesky	55.79	60	.63	–64.21	–327.02
Independent pathways	67.28	64	.37	–60.72	–341.05
Common pathways	77.92	66	.15	–54.08	–343.17
<i>Agreeableness</i>					
Cholesky	187.90	126	.00	–64.10	–616.00
Independent pathways	384.31	144	.00	96.31	–534.43
Common pathways	511.49	148	.00	215.49	–432.77
<i>Conscientiousness</i>					
Cholesky	92.89	60	.00	–27.11	–289.92
Independent pathways	108.34	64	.00	–19.66	–299.99
Common pathways	126.11	66	.00	–5.89	–294.98
<i>Openness</i>					
Cholesky	224.01	216	.34	–207.99	–1154.10
Independent pathways	452.19	256	.00	–59.81	–1181.12
Common pathways	607.75	262	.00	83.75	–1063.12

*Note.* AIC is Akaike Information Criterion; BIC is Bayesian Information Criterion; *df* is degrees of freedom.

Agreeableness in Table 7, even though the fit of the Cholesky model was superior, because the parameter estimates produced by the independent pathways model, when compared to those from the other BFM scales, highlight the nature of the difficulties with model fit for this construct.

### 3.2. Results specific to the BFM domains

#### 3.2.1. Extraversion

As described above, the common pathways model specifies that all of the covariation in a set of variables be mediated by one latent phenotypic variable influenced by a single set of additive genetic and nonshared environmental influences. Genetic influences accounted for 49% of the variance in the latent phenotype, with nonshared environmental influences accounting for the remainder. Common genetic influences (proportions of total variance) on the three adjectives ranged from .20 to .31. There were small specific genetic influences on each of the adjectives, especially “Talkative.” Specific environmental influences predominated somewhat over common influences for “Lively” and “Talkative,” with environmental influences on “Outgoing” evenly divided. The common environmental influences ranged from .21 to .33.



Table 6

Parameter estimates and variance components for the common pathways models fit to the Big Five Model adjectives for Extraversion and Neuroticism

	Factor loading	Parameter estimates		Proportion of variance accounted for by parameter			
		<i>A</i>	<i>E</i>	<i>A</i> <sup>c</sup>	<i>E</i> <sup>c</sup>	<i>A</i> <sup>s</sup>	<i>E</i> <sup>s</sup>
<i>Extraversion</i>							
Common		.70	.72	.49	.52		
Specific Outgoing	.80	.21	.57	.31	.33	.04	.32
Specific Lively	.71	.22	.67	.25	.26	.05	.43
Specific Talkative	.64	.43	.64	.20	.21	.18	.40
<i>Neuroticism</i>							
Common		.75	.66	.56	.44		
Specific Moody	.57	.35	.74	.18	.14	.12	.55
Specific Worrying	.76	.18	.63	.33	.25	.03	.40
Specific Nervous	.84	.16	.52	.40	.31	.03	.28
Specific Not calm	.45	.48	.75	.11	.09	.23	.56

*Note.* *A* is additive genetic effects in general; *E* is non-shared environmental effects in general; *A*<sup>c</sup> is common additive genetic effects; *E*<sup>c</sup> is common non-shared environmental effects; *A*<sup>s</sup> is descriptor-specific additive genetic effects; and *E*<sup>s</sup> is descriptor-specific non-shared environmental effects. Parameters are standardized.

### 3.2.2. Neuroticism

Genetic influences accounted for 56% of the variance in the latent phenotype. Common genetic influences on the four adjectives ranged from .11 to .40. There were small specific genetic influences on all four of the adjectives, particularly “Not calm.” The majority of the environmental influences were specific to individual adjectives, except for “Nervous,” for which they were evenly divided. They ranged from .09 to .56.

### 3.2.3. Agreeableness

The more structured models fit more poorly than did the Cholesky model on the basis of either AIC or BIC. Three adjectives, “Softhearted,” “Sympathetic,” and “Caring,” had lower MZ correlations, suggesting less genetic influence than on the others. At the same time, the DZ correlations varied widely, yet the variation did not appear due solely to sampling error. This led, in the independent pathways model, to estimates of no genetic influence at all, either common or specific, for “Softhearted” and “Sympathetic” and only 5% common for “Caring.” “Helpful” also received an estimate of only 7% common genetic influence, while “Friendly” and “Warm” received much higher estimates (39% and 29%, respectively). The environmental influences were split reasonably evenly between common and specific. Thus, the independent pathways model, which relies on the assumption genetic influence of some measurable magnitude across all the observed traits, could not provide as good a fit as the Cholesky, which allows for completely separate genetic influences (or ab-

Table 7

Parameter estimates and variance components for the independent pathways models fit to the Big Five Model adjectives for Agreeableness, Conscientiousness, and Openness

Factor	Parameter estimates				Proportion of variance accounted for by Parameter			
	$A^c$	$E^c$	$A^s$	$E^s$	$A^c$	$E^c$	$A^s$	$E^s$
<i>Agreeableness</i>								
Helpful	.26	.54	.35	.72	.07	.29	.12	.52
Friendly	.61	.53	.00	.59	.37	.28	.00	.35
Warm	.52	.65	.00	.55	.27	.43	.00	.30
Caring	.22	.70	.00	.68	.05	.49	.00	.46
Softhearted	.00	.57	.00	.82	.00	.33	.00	.67
Sympathetic	.00	.65	.00	.76	.00	.43	.00	.57
<i>Conscientiousness</i>								
Organized	.62	.13	.00	.77	.39	.02	.00	.60
Responsible	.38	.87	.31	.00	.14	.76	.10	.00
Hardworking	.40	.30	.31	.81	.16	.09	.09	.66
Not careless	.28	.12	.41	.86	.08	.01	.17	.74
<i>Openness</i>								
Creative	.61	.47	.00	.64	.37	.22	.00	.41
Imaginative	.54	.55	.00	.64	.29	.31	.00	.40
Intelligent	.00	.59	.39	.71	.00	.34	.15	.50
Curious	.10	.65	.19	.73	.01	.42	.04	.54
Active	.04	.62	.36	.70	.00	.39	.13	.49
Broad-minded	.03	.45	.35	.82	.00	.20	.12	.68
Sophisticated	.05	.46	.46	.76	.00	.21	.21	.58
Adventurous	.11	.61	.76	.67	.01	.37	.17	.45

Note.  $A^c$  is common genetic effects;  $E^c$  is common non-shared environmental effects;  $A^s$  is descriptor-specific additive genetic effects; and  $E^s$  is descriptor-specific non-shared environmental effects. Parameters are standardized.

sence of influences) on each observed trait. In fact, the Cholesky model included estimates of both negative and no genetic influence for paths involving “Caring,” “Softhearted,” and “Sympathetic.”

### 3.2.4. Conscientiousness

None of the models fit well, but the fit was significantly worst for the common pathways model. Allowing for model parsimony, the independent pathways model provided the best fit. The reason for the relatively poor fit the common pathways model even using BIC to place most emphasis on model parsimony was that there was little consistency among the adjectives in the extent of genetic and environmental influences on them, whether common or specific. The adjective “Not careless” showed primarily (89%) specific influences.

### 3.2.5. Openness

Considering model parsimony using BIC, the independent pathways model fit best. The reason for the relatively poor fit of the common pathways model using

BIC was that two of the eight adjectives showed strong common genetic influence while the others showed very little. The relationships between common and specific environmental influences differed among the adjectives as well.

#### 4. Discussion

In this study, we investigated the genetic and environmental structure underlying the BFM by comparing several possible latent structural models applied to data gathered on the MIDUS twin sample. To our knowledge this is the first nationwide sample of twin pairs in the United States to be recruited through screening a nationally representative sample of households. We began with a factor analysis of the personality adjectives used in the MIDUS study to represent the BFM. This factor analysis clearly revealed the BFM structure at the phenotypic level. We then fit three multivariate biometric models to the adjectives in each BFM factor domain. If there is a coherent genetic and environmental structure underlying this operationalization of the BFM, such an analysis should reveal it, and did so for Extraversion and Neuroticism when we relied on the fit statistic that places most emphasis on model parsimony. In addition, the estimates of genetic influence for the latent phenotypes of 49% and 56% for Extraversion and Neuroticism respectively were comparable to those resulting from a host of other studies of genetic influences on personality (e.g., Bouchard & Loehlin, 2001, for a review). This suggests that it is possible that Costa and McCrae's (1995) conception of these domains as latent personality constructs may be accurate.

For Agreeableness, Conscientiousness, and Openness, however, little coherent etiologic structure was revealed. It was necessary to rely upon the model fit statistic that places the most emphasis on model parsimony even to conclude that the independent pathways models provided the best fit, and this was still not possible for Agreeableness. The independent pathways model specifies direct paths from the genetic and environmental influences common to each adjective, but it also allows genetic and environmental influences specific to each adjective, so the underlying structure it imposes is rather loose. The common pathways model, which specifies a latent phenotypic personality construct for each BFM domain, fit worst for all three of these domains. For Agreeableness, the independent pathways model failed to fit better than the more general Cholesky model due to an apparent absence of either common or specific genetic influences on three of the adjectives. This suggests that, for these domains, Hofstee et al.'s (1992) conception of loose organizational categories may be accurate.

Major personality models identify constructs that are very similar to Extraversion and Neuroticism in the BFM (Block, 1965; Eysenck & Eysenck, 1985; Tellegen & Atkinson, 1974). These two constructs so pervade the literature on the structure of personality that Wiggins (1968) termed them the "Big 2." Therefore, what distinguishes the BFM is the specification of the other three domains. For Conscientiousness and Openness in this study, the independent pathways model provided the best fit when evaluated using the model fit statistic that most emphasizes model parsimony (BIC),

and the Cholesky model provided the best fit for Agreeableness. This suggests that the most parsimonious common genetic structure that can be considered to underlie the specifically BFM constructs in this study has separable and distinct effects on the propensity to endorse as characteristic the adjectives making up each construct. This implies that the genetic effects on these personality constructs derive from distinctive genetic influences acting to varying degrees on the different adjectives. Based on Hofstee et al.'s (1992) conception of the BFM as a series of circumplexes reflecting primary and secondary domain loadings, we note that, for example, within Conscientiousness, "organized" has only a primary loading, but "responsible" has a secondary loading on Agreeableness. Similarly, "creative" has only a primary loading on Openness, but "intelligent" has a secondary loading on Neuroticism as well.

At the same time, our results for Extraversion and Neuroticism make clear that it is at least possible for the kind of model comparison analysis we conducted to provide evidence that multiple genes act in concert to influence a unified personality characteristic. There are at least two possible reasons why Extraversion and Neuroticism might have produced the most coherent results in the MIDUS data. First, as mentioned above, these two domains are the most clearly defined in the sense that they are very similar to personality domains articulated in other personality models (e.g., Cloninger, 1987; Eysenck & Eysenck, 1985; Tellegen, 1982), and others have identified biological referents for them as well (e.g., Depue & Collins, 1999; Gray, 1987; Sutton & Davidson, 1997). It seems at least possible that one major reason for both the consistency with which personality researchers have specified these domains in their models and the greater coherence they present in models of the type used here is that these two domains are more etiologically "pure" in some biological sense, whether genetically or environmentally influenced. This possibility deserves greater research attention, particularly in the areas of identifying and explicating specific biological referents and mechanisms. Second, these domains had relatively few adjectives (three and four respectively), and it was thus less likely that heterogeneous sources of variance would materialize. In fact, one could question whether the three adjectives used to describe Extraversion ("Outgoing," "Lively," and "Talkative") and the four adjectives used to describe Neuroticism ("Moody," "Worrying," "Nervous," and "Not calm") could be considered to span the domains as they are usually conceptualized. We note, however, that their secondary loadings do fall on different factors in Hofstee et al.'s (1992) circumplex conception of the BFM, so lack of variation in secondary loadings for these particular adjectives does not appear to be the reason for the greater coherence for these domains. The two domains with the most adjectives, Agreeableness with six and Openness with eight, showed the least coherent results. Openness in particular showed highly heterogeneous influences. Yet number of adjectives did not entirely determine the results for the domains, as the common pathways model did not fit best for Conscientiousness, which, like Neuroticism, had four adjectives in our measure.

Our results are quite consistent with those of Jang et al. (2002). They also found large specific genetic and environmental influences on the NEO-PI-R facets in the BFM domains, and the common pathways models also failed to provide adequate fit for any of the BFM domains in their data. Their results did, however, show some-

what greater consistency for Extraversion and Neuroticism than for the other three domains. They suggested that their results called into question the meaning of the concept of a higher-order trait, but another interpretation of the results of both studies is that the difficulty in identifying unified coherent genetic and environmental influences on the five personality factors suggests problems with factor and/or construct coherence in the BFM, rather than with the concept of a latent personality trait per se. The success of the common pathways model with the admittedly limited versions of Extraversion and Neuroticism (using a fit statistic that places heavy emphasis on model parsimony) in our data provides further evidence for this possibility. We note here that our study did not make use of the NEO-PI-R in operationalizing the BFM, yet our results are quite consistent with those of Jang et al. (2002), a study that was based on the NEO-PI-R. This suggests that neither set of results is necessarily specific to the instruments used to operationalize the BFM. In addition, it is likely that our more brief and narrow measure should make it easier to support a latent phenotypic model such as the common pathways model than it would be using the more complex and heterogeneous NEO-PI-R.

The BFM was originally developed to summarize variance in lexical studies of adjectives used to describe personality across languages, and many of its proponents claim that it does no more than that, referring only to observable phenotypes at specific time points rather than to underlying causal structures such as genotypes. Yet there is substantial evidence for both high stability of personality across time and genetic influence on all commonly used measures of personality (Bouchard & Loehlin, 2001). In addition, McCrae et al. (2001) have provided evidence that the Big Five can be reproduced in genetic covariance matrices as well as in phenotypic covariance matrices. Thus the case for the BFM as a description of the organizational structure of the genetic influences on personality has been made as well. The question of the existence of an underlying coherent latent personality construct formed by each Big Five factor, however, is both more stringent and crucial to the BFM's ability to explain the ways in which genes influence personality.

The standard orthogonal (varimax) factor model typically used to extract the Big Five relies on the assumptions that the factors are uncorrelated and lie at the same second stratum (Carroll, 1993) or level in the structure imposed by the data. If, however, the biological or neurological mechanisms underlying personality at the observed stratum or level are correlated, and the manner in which they are correlated would be better described by a hierarchical structure with additional, higher level strata, then reliance on the orthogonal factor model to assess this structure will blur distinctions among personality facets and the manner in which they are organized that could be revealed by the consideration of additional possible factor models. This effect will be accentuated if the factor space spanned by the personality data underlying the analysis was designed, as were the NEO-PI-R and the adjective checklist used in this study, to reveal the BFM structure under the orthogonal factor model.

At the same time, to the extent that genes influence personality, we should expect that covariance matrices of genetic influences on personality data specifically organized to reveal the BFM under the orthogonal factor model in fact do so, as

demonstrated by McCrae et al. (2001). That is, when considered across the personality domain spanned by the data and factor analyzed in the same manner, the matrices of genetic influences may simply reveal the same structure as that imposed at the phenotypic level. If true, the BFM may exist, as many of its proponents have claimed (e.g., Goldberg, 1990), as a way in which phenotypic personality descriptors can be organized, with little or nothing to say about the etiologic structure of personality. This possibility can only be investigated by examining the underlying coherence of each domain of the Big Five as we have done here, by considering other organizational structures besides the orthogonal factor model, and by considering data organized to fit alternative models of the structure of personality. When examined in this manner, the personality models showing the most genetic coherence at the first stratum and providing the best fit to the data will have the most power to explain the etiologic structure of personality.

The results of this study have to be interpreted in light of the specific instrument used to operationalize the BFM. This instrument is both a strength, as mentioned above, and a weakness. Though the instrument was developed specifically to reveal the BFM under factor analysis and has been used successfully in other MIDUS research (e.g., Plant, Markus, & Lachman, 2002; Staudinger, Fleeson, & Baltes, 1999), in our phenotypic factor analysis of the twin sample, two of the 25 adjectives did not load most highly on the BFM factors that were intended when the BFM measure used in MIDUS was developed. In addition, some adjectives (especially “Not carelessness”) did not receive high loadings on any of the domains. This may reflect a weakness in this specific measure. On the other hand, the Big Five structure was still clearly revealed in the in the phenotypic factor analysis, and the process used to develop the measure should reveal a BFM structure if one is there to reveal. If such a structure exists and organizes etiologic influences on personality, analyses of the sort we conducted should reveal coherent etiologic structures.

In a related vein, it is possible that adjectives simply do not offer the level of measurement precision necessary to reveal the etiologic BFM structure. There are several reasons to believe, however, that this is not the case. First, Jang et al. (2002) revealed a similar lack of etiologic structure using the NEO-PI-R. Second, the standard errors around the twin correlations were small in relation both to the sizes of the correlations themselves and to most of the differences between the MZ and DZ correlations, suggesting that we were able to measure the genetic influences on the propensity to agree with each adjectival description with a reasonable level of accuracy. Third, all three models we applied provide close fits for Extraversion and Neuroticism, making it clear that this result is not impossible for a measure relying upon adjectival descriptors, at least when the number of factor descriptors is small.

It is also possible, of course, that the failure of the common pathway models for the three BFM domains beyond Extraversion and Neuroticism in this study reflects something more general about our ways of conceptualizing personality, namely the fact that many genes influence personality, and that these genes are not necessarily specific to personality traits as articulated in our natural languages. This possibility also needs to be investigated using other models of personality besides the BFM. The

ideal study would compare the performance of several models of etiological genetic and environmental structure using several different personality measurement instruments assessed in a single sample of twins. The possibility should also be investigated using Hofstee et al.'s (1992) circumplex conceptualization of the primary and secondary loadings associated with the BFM. For example, it would be interesting to examine the etiologic coherence of phenotypically pure vectors within the circumplex, and see how etiologic coherence is affected by expanding the portion of the circumplex sampled.

## Acknowledgments

This research was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development, and by National Institute on Aging Grant # AG20166. We thank Tom Bouchard, Kerry Jang, John Loehlin, Kristian Markon, Matt McGue, Chris Patrick, Carol Ryff, Auke Tellegen, and David Watson for their helpful comments on an earlier draft of the manuscript.

## References

- Akaike, H. (1983). Information measures and model selection. *Bulletin of the International Statistical Institute*, 50, 277–290.
- Block, J. (1965). *The challenge of response sets*. New York: Appleton-Century-Crofts.
- Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2002). Similarity of childhood experiences and personality resemblance in monozygotic and dizygotic twins: A test of the equal environments assumption. *Personality and Individual Differences*, 33, 261–269.
- Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, 31(3), 243–273.
- Bouchard, T. J., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*, 54, 4–45.
- Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2001). Comprehensive Exploratory Factor Analysis 1.03b.
- Carroll, J. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.
- Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants. *Archives of General Psychiatry*, 44, 573–588.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64, 21–50.
- Costa, P. T., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from Conscientiousness. *European Journal of Personality*, 12, 117–134.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, 22, 491–569.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. New York: Plenum.
- Finkel, D., & McGue, M. (1997). Sex differences and nonadditivity in heritability of the Multidimensional Personality Questionnaire Scales. *Journal of Personality and Social Psychology*, 72(4), 929–938.
- Goldberg, L. R. (1990). An alternative “Description of personality”: The Big-Five structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Goldberg, L. R. (1992). The development of markers for the Big-Five structure. *Psychological Assessment*, 4, 26–42.

- Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge, England: University of Cambridge Press.
- Heath, A. C., Eaves, L. J., & Martin, N. G. (1989). The genetic structure of personality: III. Multivariate genetic item analysis of the EPQ scales. *Personality and Individual Differences*, *10*, 877–888.
- Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*, 146–163.
- Jang, K. L., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, *33*, 83–101.
- 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 language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.
- Johnson, W., Krueger, R. F., Bouchard, T. J., & McGue, M. (2002). The personalities of twins: Just ordinary folks. *Twin Research*, *5*(2), 125–131.
- Kendler, K. S., Thornton, L. M., Gilman, S. E., & Kessler, R. C. (2000). Sexual orientation in a US national sample of twin and sibling pairs. *American Journal of Psychiatry*, *157*(11), 1843–1846.
- Kessler, R. C., Gilman, S. E., Thornton, L. M., Kendler, K. S. (in press). Health, wellbeing, and social responsibility in the MIDUS twin and sibling subsamples. In O. G. Brim, C. D. Ryff, & R. C. Kessler (Eds.), *How healthy are we: A national study of wellbeing at midlife*. Chicago: University of Chicago Press.
- Krueger, R. F., Markon, K. M., & Bouchard, T. J. (2003). The extended genotype: The heritability of personality accounts for the heritability of recalled family environments in twins reared apart. *Journal of Personality*, *71*, 809–833.
- Lachman, M. E., & Weaver, S. L. (1997). The Midlife Development Inventory (MIDI) personality scales: Scale construction and scoring. Waltham, MA.
- Loehlin, J. C., & Nichols, J. (1976). *Heredity, environment and personality*. Austin: University of Texas.
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Newberry Park, CA: Sage.
- Loehlin, J. C., McCrae, R. R., Costa, P. T., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, *32*, 431–453.
- Lykken, D. T., Bouchard, T. J., McGue, M., & Tellegen, A. (1990). The Minnesota Twin Family Registry: Some initial findings. *Acta Genetica Medica Gemellol*, *39*, 35–70.
- McCrae, R. R., & Costa, P. T. (1990). *Personality in adulthood*. New York: Guilford.
- McCrae, R. R., & Costa, P. T. (1999). A five-factor model of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 139–153). New York: The Guilford Press.
- McCrae, R. R., Jang, K. L., Livesley, W. J., Riemann, R., & Angleitner, A. (2001). Sources of structure: Genetic, environmental, and artifactual influences on the covariation of personality traits. *Journal of Personality*, *69*(4), 511–535.
- McGue, M., & Bouchard, T. J., Jr. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics*, *14*, 325–343.
- Neale, M. C. (1997). *Mx: Statistical modeling*. Richmond, VA: Medical College of Virginia, Department of Psychiatry.
- Paunonen, S. V., & Jackson, D. N. (1996). The Jackson Personality Inventory and the Big Five Model of personality. *Journal of Research in Personality*, *30*, 42–59.
- Plant, V. C., Markus, H. R., & Lachman, M. E. (2002). Place matters: Consensual features and regional variation in American well-being and self. *Journal of Personality and Social Psychology*, *83*, 160–184.
- Priest, R. A., & Vandenberg, S. G. (1980). Spouse similarity in American and Swedish couples. *Behavior Genetics*, *10*, 59–71.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.
- 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. (1996a). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, *10*, 61–77.
- Saucier, G., & Goldberg, L. R. (1996b). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Perspectives* (pp. 21–50). New York: Guilford.
- Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, *76*, 613–627.
- Staudinger, U. M., Fleeson, W., & Baltes, P. B. (1999). Predictors of subjective physical health and global well-being: Similarities and differences between the United States and Germany. *Journal of Personality and Social Psychology*, *76*, 305–319.
- Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, *8*, 204–210.
- Tellegen, A. (1982). *Brief manual for the differential personality questionnaire*. Minneapolis.
- Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences (“Absorption”), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology*, *83*, 268–277.
- Tellegen, A., Bouchard, T. J., Wilcox, K. J., Segal, N. L., Lykken, D. T., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology*, *54*(6), 1031–1039.
- Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology*, *59*, 781–790.
- Wiggins, J. S. (1968). Personality structure. *Annual Review of Psychology*, *19*, 293–350.
- Waller, N. G. (1999). Evaluating the structure of personality. In C. R. Cloninger (Ed.), *Personality and psychopathology*. Washington, DC: American Psychiatric Association.