The Big Five Personality Traits: Psychological Entities or Statistical Constructs?

Article *in* Behavior Genetics · October 2013

DOI: 10.1007/s10519-013-9625-7 · Source: PubMed

4 authors, including:

Denny Borsboom
University of Amsterdam
227 PUBLICATIONS **11,480** CITATIONS

SEE PROFILE

Conor Dolan
University of Amsterdam
252 PUBLICATIONS **6,853** CITATIONS

SEE PROFILE

Dorret I Boomsma
Vrije Universiteit Amsterdam
2,070 PUBLICATIONS **77,633** CITATIONS

SEE PROFILE

Behav Genet
DOI 10.1007/s10519-013-9625-7

ORIGINAL RESEARCH

The Big Five Personality Traits: Psychological Entities or Statistical Constructs?

Sanja Franić · Denny Borsboom · Conor V. Dolan · Dorret I. Boomsma

Received: 29 June 2013 / Accepted: 14 October 2013

Abstract The present study employed multivariate genetic item-level analyses to examine the ontology and the genetic and environmental etiology of the Big Five personality dimensions, as measured by the NEO Five Factor Inventory (NEO-FFI) [Costa and McCrae, Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI) professional manual, 1992; Hoekstra et al., NEO personality questionnaires NEO-PI-R, NEO-FFI: manual, 1996]. Common and independent pathway model comparison was used to test whether the five personality dimensions fully mediate the genetic and environmental effects on the items, as would be expected under the realist interpretation of the Big Five. In addition, the dimensionalities of the latent genetic and environmental structures were examined. Item scores of a population-based sample of 7,900 adult twins (including 2,805 complete twin pairs; 1,528 MZ and 1,277 DZ) on the Dutch version of the NEO-FFI were analyzed. Although both the genetic and the environmental covariance components display a 5-factor structure, applications of common and independent pathway modeling showed that they do not comply with the collinearity constraints entailed in the common pathway model. Implications for the substantive interpretation of the Big Five are discussed.

S. Franić (✉) · C. V. Dolan · D. I. Boomsma
Department of Biological Psychology, Faculty of Psychology and Education, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
e-mail: s.franic@vu.nl

D. Borsboom
Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands

Keywords Personality · Big Five · NEO-FFI · Genetics · Item-level analysis · Common pathway model · Independent pathway model

Introduction

Over the past century, one of the most influential approaches to personality description has been the five factor (FF) approach. Predicated on the lexical approach to personality description, reflected in Cattell (1943a): "All aspects of human personality which are or have been of importance, interest, or utility have already become recorded in the substance of language" (p. 483), the FF approach is based on the idea that identification of basic dimensions of human personality is possible via the application of factor analytic techniques to verbal descriptors of human traits.

The beginnings of the FF approach can be traced to Allport and Odbert (1936) selection of 4,504 psychological trait terms from the 1925 unabridged Webster's New International Dictionary. Cattell (1943a, b, 1945) augmented this list in the 1940s by adding "the substance of all syndromes and types which psychologists have observed and described in the past century or so", and subsequently abbreviated it to a set of 35 variables—a factor analysis of which produced 12 "primary" factors. In the early 1960s, Tupes and Christal (1992) performed a series of factor analyses on Cattell's variables and observed five recurrent orthogonal factors, which they denoted surgency/extraversion, agreeableness, dependability, emotional stability, and culture (French 1953). Through Norman's (1963, 1967) further addition to, and subsequent abbreviation of, Allport and Odbert's original list, and further selection of adjectives from this list by Goldberg (1977, 1980, 1983, 1990, 1992), a set of variables with a clearer five-factor

Published online: 27 October 2013

orthogonal structure was produced. Goldberg (1981, 1982) denoted these FFs "the Big Five".

In a parallel research program, following a cluster analysis of Cattell's Sixteen Personality Factor (16PF) Questionnaire in which three factors were extracted—neuroticism, extraversion, and openness to experience—McCrae and Costa (1983) developed a 144-item, 18-facet, 3-dimensional questionnaire, which they termed the NEO Inventory. After linking their neuroticism and extraversion factors to those from the previous lexically based research (e.g., Goldberg 1980, 1981, 1982, 1983), they fully adopted the FF approach, and consequently developed measures of agreeableness and conscientiousness. The addition of these scales to the NEO Inventory resulted in the NEO Personality Inventory (Costa and McCrae 1985), and the subsequent implementation of facets to measure these two new factors yielded the Revised NEO Personality Inventory (NEO-PI-R; Costa and McCrae 1992). The NEO Five Factor Inventory (NEO-FFI) is a shorter, 60-item version of the NEO-PI-R.

The FF approach has been extraordinarily influential: numerous behavior genetics studies have assessed the heritabilities of the Big Five (and more recently sought associations with measured genetic variants; de Moor et al. 2010), neural and clinical correlates of the FFs have been examined (e.g., DeYoung et al. 2010; Nigg et al. 2002), and the model has found wide practical application, for instance in the field of personnel selection (Schmit and Ryan 1993). A google scholar search for "FF model personality" returns nearly two million hits, and a google search of the same term returns around 121 million.

Notwithstanding its popularity, however a plethora of issues have been raised concerning the conceptual, empirical and statistical foundations of the FF approach (e.g., Block 1995). Lack of formal theory underpinning the approach and the possibility of empirical analyses being shaped by prior conceptual commitments are some of the most prominent ones. Concerns have been raised over the orthogonality of the factor solutions, their proposed simple structure, and even the number of factors being significantly impacted by the selection of input variables and choices of factor rotations, which ultimately might have rested more on the authors' conceptual beliefs than on mathematical/statistical criteria. In addition, the degree of arbitrariness involved in Cattell, Norman, and Goldberg's selection of trait terms and construction of clusters remains unknown. The model has received additional criticism for failing to account for intra-individual personality structure and personality functioning. A factor analysis of common english terms describing laptop computers, for instance, might yield size, processing speed, random-access memory capacity, storage capacity, and operating system as five orthogonal factors; however, one may wonder to what extent these factors are informative about the physical structure of a laptop computer, or its functional architecture (Cervone 2005). The model has also been criticized on psychometric grounds for a number of problems including failure of orthogonality (Costa and McCrae 1992; Block and Block 1980; Mroczek 1992; Goldberg 1992, 1993; Peabody and Goldberg 1989), the presence of cross-loadings (Parker et al. 1993; Block 1995; Costa and McCrae 2008), low validity coefficients (Pervin 1994), lack of reproducibility of the five-factor structure from other personality inventories (Caprara et al. 1995; Hahn and Comrey 1994), and lack of fit in confirmatory context (Parker et al. 1993; McCrae et al. 1996). The FF model is derived through, and based on, exploratory techniques such as exploratory factor analysis (EFA) and principal components analysis (PCA); in the confirmatory factor analysis (CFA) context, however, the model typically obtains unsatisfactory fit.

Another, arguably more fundamental issue, concerns a possible misinterpretation of principal components (Markus and Borsboom 2013) and, more broadly, the ontological nature of the FFs. Being generated in a formative model, the components obtained in PCA are efficient statistical summaries of the data. Their standard interpretation amongst FF model proponents, however, is of a realist nature; they are considered to be behavior-generating entities (e.g., extraversion causes party-going behavior; McCrae and Costa 2008). This possible misinterpretation of principal components, along with some of the other criticism listed above, has prompted questions about whether the Big FFs are truly a discovery, as advocated by its proponents, or should rather be seen as a set of statistical constructs emanating from factor analysis of possibly preselected sets of variables.

In the present article, we address the last issue using quantitative genetic methodology. As outlined in a recent article by Franić et al. (2013a), quantitative genetic methods can be used to test hypotheses regarding the ontological nature of latent variables. In particular, we address the question of whether the realist interpretation of the Big Five personality factors (in which the factors represent entities that cause the observed item responses) is supported by the data, or whether the factors would more correctly be interpreted as statistical constructs. To this end, we examine the dimensionality of the latent genetic and environmental structures underlying the observed covariation in NEO-FFI items. Behavior genetic studies have been performed on personality data before (e.g., Loehlin 1989; Loehlin and Martin 2001; Bouchard and Loehlin 2001; Plomin and Caspi 1990), but item-level analyses, which make it possible to address the specific research question above, have seldom been undertaken on NEO-FFI or NEO-PI-R data (but see Johnson and Krueger 2004).

Approach

Genetic covariance structure modeling (Martin and Eaves 1977) is the application of structural equation modeling (Bollen 1989; Kline 2005) to data collected in genetically informative samples, such as samples of twins (Neale and Cardon 1992; Franić et al. 2012). In the classical twin design, the sample consists of monozygotic (MZ) and dizygotic (DZ) twin pairs. DZ twins share an average of 50 % of their segregating genes, while MZ twins share nearly their entire genome (Falconer and Mackay 1996; van Dongen et al. 2012). The covariance structure of the phenotypes (i.e., observed traits) is modeled as a function of latent factors representing several sources of individual differences: additive genetic (A), non-additive genetic (D), shared environmental (C), and individual-specific environmental (E) sources. Additive genetic influences are modeled by one or more A factors, which represent the total additive effects of genes relevant to the phenotype. Non-additive genetic influences are modeled by one or more D factors, representing the total non-additive effects of genes relevant to the phenotype. Non-additive effects arise from interactions of alleles within the same locus (genetic dominance) and/or across different loci (epistasis). Based on quantitative genetic theory (Falconer and Mackay 1996; Mather and Jinks 1971), the A factors are known to correlate one across MZ twins and five across DZ twins, and D factors are known to correlate one across MZ twins and 25 across DZ twins. Environmental influences affecting the phenotype of both twins in an identical way, thereby increasing their similarity beyond what is expected based on genetic resemblance alone, are represented by one or more C factors. Therefore, by definition, the C factors correlate unity across twins (regardless of zygosity). All environmental influences causing the phenotype of two family members to differ are represented by one or more E factors. Thus, by definition, the E factors are uncorrelated across twins.

The classical twin design does not allow for simultaneous estimation of A, C and D effects (Keller and Coventry 2005); two of these sources of individual differences can be modeled at most.[1] Assuming, for instance, an ADE model, the expected covariance structure in a multivariate twin model is:

$$\begin{matrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{matrix} = \begin{matrix} \boldsymbol{\Sigma}_{A} + \boldsymbol{\Sigma}_{D} + \boldsymbol{\Sigma}_{E} & r_{A}\boldsymbol{\Sigma}_{A} + r_{D}\boldsymbol{\Sigma}_{D} \\ (r_{A}\boldsymbol{\Sigma}_{A} + r_{D}\boldsymbol{\Sigma}_{D})^{t} & \boldsymbol{\Sigma}_{A} + \boldsymbol{\Sigma}_{D} + \boldsymbol{\Sigma}_{E}, \end{matrix} \quad (1)$$

where given p phenotypes, $\boldsymbol{\Sigma}_{11}$ ($\boldsymbol{\Sigma}_{22}$) is the p × p covariance matrix of twin 1 (twin 2), $\boldsymbol{\Sigma}_{12}$ ($\boldsymbol{\Sigma}_{21}$) is the twin 1–twin 2 p × p covariance matrix, and $\boldsymbol{\Sigma}_{A}$, $\boldsymbol{\Sigma}_{D}$ and $\boldsymbol{\Sigma}_{E}$ are the additive genetic, non-additive genetic, and unique environmental p × p covariance matrices, respectively. The coefficients r_A and r_D are the additive and the non-additive genetic twin correlations, respectively (MZ: $r_A = r_D = 1$; DZ: $r_A = 1/2$, $r_D = 1/4$).

Figure 1 gives two examples of the multivariate twin models used in the present study. The first model in Fig. 1 is a common pathway model (Kendler et al. 1987), also known as the psychometric factor model (McArdle and Goldsmith 1990). In a common pathway model, all of the A,C(D), and E influences on the item responses are mediated by a latent variable, also referred to as the psychometric factor (factors P_1 and P_2 in Fig. 1). P_1 and P_2 may be viewed as latent factors obtained in standard psychological research, e.g. 'neuroticism' or '*g*'. The second model in Fig. 1 is an independent pathway model (Kendler et al. 1987), also known as the biometric factor model (McArdle and Goldsmith 1990). In the independent pathway model, there is no phenotypic latent variable that mediates genetic and environmental effects on the item responses. Rather, the A, C(D), and E factors influence item responses directly.

In the present text, we distinguish between genetic factor models (introduced above), and phenotypic factor models. By 'phenotypic factor model', we refer to the factor model as usually formulated and applied in psychological research. The term 'phenotypic' is used to indicate that the model is applied to observed (i.e., phenotypic) covariation; no genetic information is used. The 8-factor cross-informant model of the CBCL (Achenbach 1991) and the FF model of personality (McCrae and Costa 1999; McCrae and John 1992) are examples of a phenotypic factor model.

The common pathway model bears a number of similarities to the phenotypic factor model. Notably, both the phenotypic factor model and the common pathway model are based on the premise that all covariation in item responses is attributable to one or more latent variables. In phenotypic factor modeling, this hypothesis can be formulated in terms of measurement invariance: all external variables that produce covariation in item responses exert their influence via the latent variable (Mellenbergh 1989; Meredith 1993). Likewise, in common pathway modeling, one assumes that all of the A, C(D), and E effects on item covariation are mediated by the psychometric factor. That is, there are no direct effects of A, C(D), and E on the items.

The assumption of full mediation of external influences by the latent phenotypic variable(s) has strong implications. For instance, different external variables affecting a set of item responses via the same latent variable exert the same magnitude of influence relative to each other on all the items that depend on that latent variable. For instance,

[1] Other designs, e.g., the nuclear twin family design, the stealth design, or the cascade design permit simultaneous estimation of A, C, D and E effects (Keller et al. 2010).

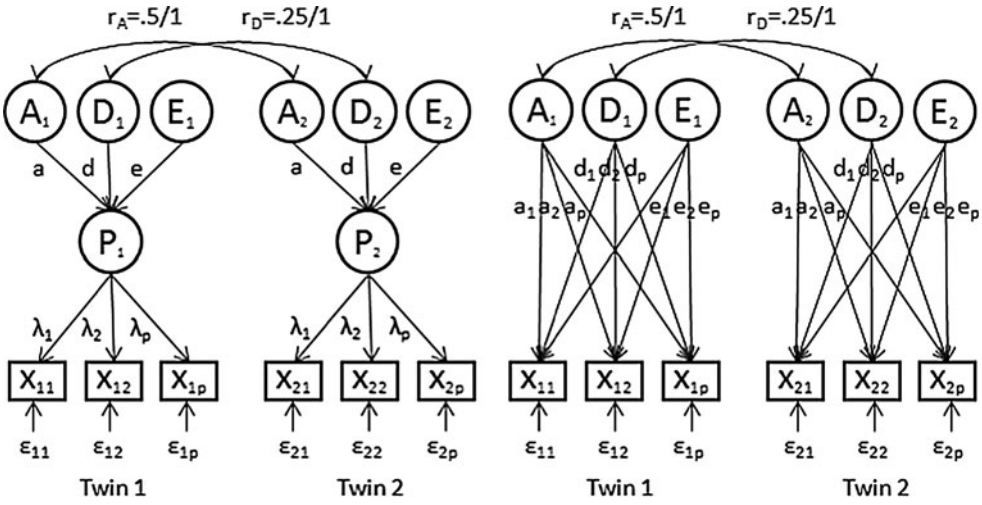

Fig. 1 A common (*left*) and an independent (*right*) pathway model

if an A and an E variable affect a set of items via the same psychometric factor, then the magnitude of influence exerted by the variable A on any individual item will be a scalar multiple of the magnitude of influence exerted by the variable E on the same item, and this scalar multiple (k) will be a constant across all the items depending on the same psychometric factor. This means that one can derive a common pathway model from an independent pathway model by imposing proportionality constraints on the factor loadings, such that $a_1/a_2 = d_1/d_2 = e_1/e_2 = k$ (following the notation in the right panel of Fig. 1).

Thus, the common pathway model makes explicit an assumption of the phenotypic latent variable model concerning the sources of item covariation: all influences on item covariation are mediated by the phenotypic latent variable. Barring exceptional cases of model equivalence, this means that a latent variable model cannot hold unless the corresponding common pathway model holds (Franić et al. 2013a). Because any given latent variable hypothesis implies a corresponding common pathway model, a refutation of that common pathway model would constitute evidence against the latent variable hypothesis.

For this reason, one may test the latent variable hypothesis by comparing the fit of a common pathway model to that of a corresponding independent pathway model. Specifically, if a model in which all of the A, C(D), and E factors exert direct influence on the phenotype fits the data statistically better than a model in which these influences are mediated by a phenotypic latent variable, this provides evidence against the hypothesis that the effects on the observed item covariation are completely mediated by the phenotypic latent variable. In that case, the latent factors employed in the phenotypic factor model are no more than an amalgamation of the direct influences of the A, C(D), and E factors on the observed item responses. If, on the other hand, an independent pathway model does not fit the data better than the corresponding common pathway model, this would provide support for the structure employed in the common pathway model, and substantiation for the corresponding phenotypic latent variable hypothesis. Comparison of an independent pathway model and a common pathway model may be conducted using a likelihood ratio test, because, as shown, a common pathway model can be derived from an independent pathway model by imposing appropriate proportionality constraints on the factor loadings (i.e., the models are nested).

Methods

Data

The data were obtained from the Netherlands Twin Register at VU University Amsterdam (Willemsen et al. 2013), and consist of item scores of a population-based sample of 7,900 adult twins (including 2,805 complete twin pairs; 1,528 MZ and 1,277 DZ) on the Dutch version of the NEO-FFI (Costa and McCrae 1992; Hoekstra et al. 1996). The participants were aged between 18 and 86 ($M = 32.3$, $SD = 12.7$) at time of measurement. 68.3 % of the participants were female. The NEO-FFI is a 60-item personality questionnaire consisting of 5 subscales: Neuroticism (N), extraversion (E), openness (O), agreeableness (A), and conscientiousness (C). Item content is given in Table 1. The responses are given on a 5-point scale ('strongly disagree', 'disagree', 'neutral', 'agree', 'strongly agree').

Initially, the sample consisted of 8,090 twins, and missingness was limited to 0.9 %. In treating missingness, we adopted the guidelines outlined in the NEO-FFI manual (Costa and McCrae 1992; Hoekstra et al. 1996): if missingness per participant exceeded 15 %, the participant's scores were excluded from the analyses. The application of this criterion reduced the missingness to 0.4 %, and the

Table 1 Item content of the NEO-FFI. Item numbering in the parentheses corresponds to that in the text/Tables/Figures

Item no.		Item content	Scale
1	(n1)	Not a worrier*	Neuroticism
6	(n2)	Feels inferior	
11	(n3)	Goes to pieces under stress	
16	(n4)	Rarely lonely or blue*	
21	(n5)	Tense, jittery	
26	(n6)	Feels worthless	
31	(n7)	Rarely fearful or anxious*	
36	(n8)	Angry at the way people treat him	
41	(n9)	Easily discouraged	
46	(n10)	Seldom sad or depressed*	
51	(n11)	Feels helpless	
56	(n12)	Ashamed	
2	(e1)	Likes having many people around	Extraversion
7	(e2)	Laughs easily	
12	(e3)	Not cheerful or light-hearted*	
17	(e4)	Enjoys talking to people	
22	(e5)	Likes to be where the action is	
27	(e6)	Prefers to do things alone*	
32	(e7)	Bursting with energy	
37	(e8)	Cheerful, vivacious	
42	(e9)	Not a cheerful optimist*	
47	(e10)	Leads a fast-paced life	
52	(e11)	Very active	
57	(e12)	Rather go his own way than lead others*	
3	(o1)	Doesn't waste time daydreaming*	Openness
8	(o2)	Sticks to a single way of doing things*	
13	(o3)	Intrigued by patterns	
18	(o4)	Thinks controversial speakers only confuse students*	
23	(o5)	Not affected by poetry*	
28	(o6)	Tries new foods	
33	(o7)	Doesn't notice moods different environments produce*	
38	(o8)	Looks to religious authorities for moral decisions*	
43	(o9)	Excited by poetry or art	
48	(o10)	Little interest in speculating about nature of universe*	
53	(o11)	Wide range of intellectual interests	
58	(o12)	Enjoys playing with theories	
4	(a1)	Courteous	Agreeableness
9	(a2)	Often gets into arguments*	
14	(a3)	Some consider him selfish or egotistical*	
19	(a4)	Prefers cooperation to competition	
24	(a5)	Cynical, skeptical*	
29	(a6)	Thinks people will take advantage*	
34	(a7)	Most people like him	
39	(a8)	Some consider him cold or calculating*	
44	(a9)	Business-like, unsentimental*	
49	(a10)	Thoughtful, considerate	
54	(a11)	Shows if he doesn't like people*	
59	(a12)	Prepared to manipulate*	

Table 1 continued

Item no.		Item content	Scale
5	(c1)	Keeps belongings neat and clean	Conscientiousness
10	(c2)	Good at pacing himself	
15	(c3)	Not very methodical*	
20	(c4)	Performs tasks conscientiously	
25	(c5)	Has a clear set of goals	
30	(c6)	Wastes time before settling down to work*	
35	(c7)	Works hard	
40	(c8)	Follows through on commitments	
45	(c9)	Not dependable*	
50	(c10)	Productive	
55	(c11)	Unable to get organized*	
60	(c12)	Strives for excellence	

Reverse-scored items are marked with an *

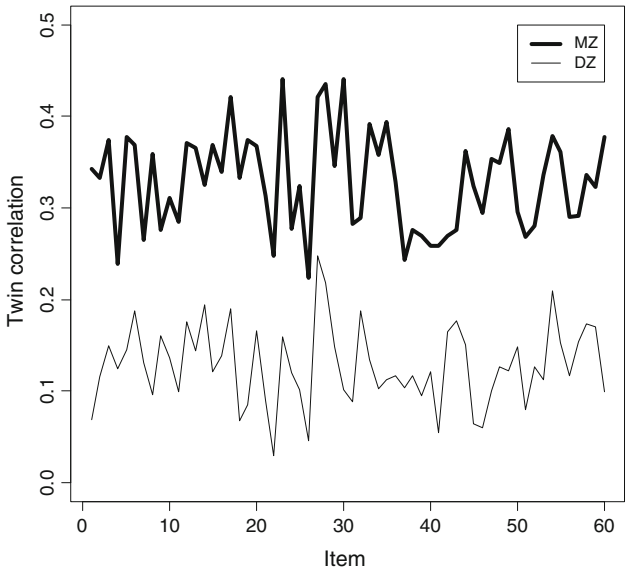

Fig. 2 Phenotypic polychoric MZ and DZ twin item correlations

sample size to $N = 7{,}900$. The remaining missing values were assigned the 'neutral' value of 3. Application of LISREL's (Jöreskog and Sörbom 2004) test for underlying bivariate normality indicated no significant departures from normality for any of the items. The MZ and DZ twin item correlations are depicted in Fig. 2.

Analyses

In the first phase of the analyses, the phenotypic structure of the NEO-FFI was examined using EFA and CFA. Here, the data were treated as if the sample consisted of unrelated individuals. To correct for the clustering in the data due to the genetic relatedness, we employed a correction for clustering available in MPlus (Rebollo et al. 2006; Muthén and Muthén 1998–2007). EFA and CFA were performed using split-half validation: EFA was performed on one randomly selected half of the sample ($N = 3{,}950$), and CFA on the other ($N = 3{,}950$). In EFA, 3–6 factor solutions with the oblique geomin rotation were tested. We opted for an oblique criterion because the NEO-PI-R and NEO-FFI data conform appreciably better to a model with oblique factors, despite the initial idea of orthogonality (Mroczek 1992; Goldberg 1993; Peabody and Goldberg 1989). The best-fitting substantively interpretable model indicated by EFA was subsequently tested in CFA.

In the second phase of the analyses, the results of the phenotypic analyses were used as a basis for specifying multivariate common and independent pathway genetic factor models. Here, only data on complete twin pairs (1,528 MZ and 1,277 DZ twin pairs) were used. The genetic and environmental etiology of the items was first examined using univariate modeling: a number of competing models (ACE, ADE, AE) were fitted to each of the 60 items, and likelihood ratio testing was employed to determine the best model for each item. The same approach was used on subscale level: univariate (ACE, ADE, AE) models were fitted to each of the five subscales. The results of these preliminary analyses were subsequently used as a basis for specifying the common and independent pathway models.[2]

To address the central question concerning the ontological nature of the latent personality factors, the common and independent pathway models were compared against each other using likelihood ratio testing. Finally, to explore the structure of the genetic and environmental influences

[2] Although item-specific residual factors can be subjected to their own AC(D)E decomposition, in the present paper this was not done given our focus on dimensionality assessment and the common/independent pathway model comparison. The residual covariances between the twins were however added. These covariances were estimated separately in the MZs and DZs, given the possible genetic residual effects.

on the NEO-FFI items in a hypothesis-free fashion, the 60 × 60 phenotypic covariance matrix was decomposed into 60 × 60 genetic and environmental correlation matrices, and each of these matrices was subjected to EFA. The genetic and environmental correlations matrices were obtained in a standard twin model using Cholesky decompositions in Mx (Neale 2000). We used the phenotypic 120 × 120 (60 per twin) polychoric correlation matrix as input, because Pearson product moment correlations based on discrete data tend to be slightly biased (Dolan 1994).

The analyses were carried out using Mplus 5 (Muthén and Muthén 1998–2007), Mx, and R (R Development Core Team 2009). Given the discrete nature of the items, we fitted discrete factor models (i.e., we assumed the discrete indicator variables to be a realization of a continuous normal latent process, and modeled polychoric correlations; Flora and Curran 2004) using the robust weighted least squares estimator (WLSMV; Muthén and Muthén 1998–2007). The polychoric correlations between the 60 items and between the 120 (60 per twin) items served as input in the phenotypic and the genetic factor analyses, respectively. In evaluating model fit, the Tucker Lewis Index (TLI)[3] and the Root Mean Square Error of Approximation (RMSEA) were used. Cut-off values of >0.90 TLI and <0.08 RMSEA were employed as criteria for acceptable fit. As both our sample size and the models employed were large, the Chi square statistic was of limited use as an overall fit measure (Jöreskog 1993), and was employed only to test local hypotheses concerning comparisons of nested models, as these comparisons are associated with a smaller approximation error.

Results

The results of the phenotypic EFA are given in Tables 2 and 3. As evident from Table 2, the 5- and the 6-factor phenotypic solution both fitted adequately (TLI >0.94, RMSEA <0.055). However, as the 6-factor solution was difficult to interpret substantively, in further analyses we focused on the 5-factor solution. This solution, detailed in Table 3, resembles closely Costa and McCrae's (1992) FF model.

Based on the EFA results, a 5-factor model [corresponding exactly to Costa and McCrae's (1992) FF model] was formulated and tested in CFA. The fit measures, given in Table 2, indicated a suboptimal fit. This is not unexpected considering the literature, which frequently reports a misfit of the FF model to empirical data (e.g., Parker et al. 1993; McCrae et al. 1996). To examine the extent to which the misfit is due to presence of cross-loadings, in the next step we freed all the cross-loadings with a modification index larger than 50, and re-fitted the model. This resulted in an acceptable model fit ($\chi^2 = 9{,}708$, df = 499, TLI = 0.899, RMSEA = 0.068). However, the modified model contained 94 cross-loadings.

Table 2 Fit measures for the 3–6-factor geomin-rotated EFA solutions and the 5-factor CFA model

Method	Factors	χ^2	df	TLI	RMSEA
EFA	3f	19,932	535	.822	.094
	4f	13,918	599	.887	.075
	5f	8,266	648	.940	.055
	6f	6,827	642	.951	.049
CFA	5f	17,222	436	.789	.099

Table 4 shows the factor loadings, residual variances and inter-factor correlations associated with the simple structure 5-factor model. The average variance explained by the factors ranges from 22 % (O and A factors) to 42 % (N factor). The factor correlations between O and the other factors are generally low ($r < 0.12$). The correlations between N and the remaining factors are substantial and negative (from −0.41 to −0.62), and the rest of the factors (E, A, and C) are substantially and positively intercorrelated; from 0.45 to 0.48. This is line with the literature, which frequently reports substantial correlations between the FFs (e.g., Block 1995).

In the first step of the genotypic analyses, the genetic and environmental etiology of the items was examined in a univariate fashion. The same was done on the subscale level, with the subscale scores being defined as the sum scores across the relevant items. Overall, none of the items or scales contained a detectable C component. With regard to the A, D, and E influences, the items displayed two major patterns: some appeared additive genetic and unique environmental in origin (AE model), while for the rest neither additive nor non-additive genetic influences could be detected (E model). On subscale level, only the Agreeableness scale displayed a significant D component, and the remaining scales conformed to an AE model. As another set of our preliminary analyses showed that the D component did not exceed 5 % for any of the items (M = 2.1 %, SD = 1 %), and that a D component was only detected for a limited number of items, D was not modeled in the subsequent analyses. Considering that the power to detect sources of variation is greater at the

[3] TLI is an incremental fit index based on the difference in fit of a baseline model with uncorrelated variables and the fitted model. The standard rule of thumb was formulated for the analyses of scale scores, not item score. As item scores tend to correlate to a lesser extent than scale scores (often based on multiple items), the standard TLI rule of thumb is hard to satisfy. See e.g. Kenny (2012).

Table 3 Standardized factor loadings (λ), residual variances (μ_R = mean residual variance) and inter-factor correlations in the geomin–rotated 5-factor phenotypic EFA solution

Item	λ_N	λ_E	λ_O	λ_C	λ_A	Residual variance
n1	**0.55**				0.15	0.67
n2	**0.63**					0.55
n3	**0.65**				0.13	0.56
n4	**0.57**	−0.15				0.61
n5	**0.76**					0.43
n6	**0.75**			−0.10		0.37
n7	**0.66**				0.18	0.54
n8	**0.56**				−0.28	0.61
n9	**0.57**		−0.11	−0.28		0.49
n10	**0.67**	−0.12			0.18	0.49
n11	**0.59**			−0.27		0.48
n12	**0.56**			−0.13		0.64
						$\mu_R = 0.54$
e1		**0.67**	−0.11			0.58
e2	−0.18	**0.65**				0.49
e3	−0.45	**0.53**			0.11	0.39
e4		**0.63**			0.13	0.53
e5		**0.61**			−0.21	0.61
e6	−0.19	**0.32**		−0.15	0.21	0.80
e7	−0.22	**0.34**		0.21	−0.26	0.64
e8	−0.31	**0.71**				0.26
e9	−0.43	**0.44**				0.50
e10	.	0.23		0.14	**−0.29**	0.83
e11	−0.15	**0.41**		0.41	−0.19	0.48
e12	**−0.26**	0.14	0.11		−0.14	0.86
						$\mu_R = 0.58$
o1	0.14		0.26	**−0.32**	0.12	0.79
o2	−0.11	−0.13	**0.22**	−0.20		0.91
o3		−0.11	**0.60**			0.64
o4	−0.19		**0.36**		0.16	0.82
o5			**0.62**		0.16	0.60
o6		0.15	**0.26**			0.90
o7			**0.31**	0.10	0.25	0.81
o8	−0.11		**0.15**			0.96
o9			**0.67**			0.53
o10			**0.53**			0.71
o11	−0.13	0.14	**0.53**	0.17	−0.14	0.59
o12			**0.58**		−0.21	0.63
						$\mu_R = 0.74$
a1	0.13	0.18		**0.27**	0.25	0.80
a2	−0.27	−0.13			**0.43**	0.71
a3				0.16	**0.53**	0.63
a4		0.13			**0.38**	0.83
a5	−0.37				**0.44**	0.62
a6	**−0.43**		0.12		0.34	0.70
a7	−0.13	**0.36**		0.20	0.20	0.68
a8		0.11			**0.63**	0.58
a9		0.20			**0.58**	0.63
a10	0.21	0.29		0.39	**0.40**	0.55

Table 3 continued

Item	λ_N	λ_E	λ_O	λ_C	λ_A	Residual variance
a11					**0.42**	0.81
a12			−0.16	0.12	**0.54**	0.65
						$\mu_R = 0.68$
c1		−0.14	−0.11	**0.57**	0.14	0.66
c2	−0.16			**0.62**		0.55
c3		−0.17		**0.49**		0.74
c4	0.10			**0.45**	0.29	0.70
c5				**0.62**	−0.24	0.56
c6	−0.19	−0.18	−0.15	**0.55**		0.58
c7		0.19		**0.67**	−0.10	0.49
c8				**0.56**	0.20	0.62
c9	−0.14		−.12	**0.41**	0.29	0.66
c10		0.12		**0.61**	−0.12	0.53
c11	−0.46			**0.46**	0.11	0.44
c12	0.16	0.19	0.13	**0.38**	−0.27	0.72
						$\mu_R = 0.60$
Factor correlations:						
	N	E	O	A		
E	−0.25					
O	0.08	0.13				
A	−0.31	0.24	0.07			
C	−0.06	0.03	0.00	0.09		

The highest loading for each item is given in bold. Factor loadings smaller than 0.10 are omitted

subscale level, which conformed predominantly to an AE model, we proceeded with the multivariate analyses using an AE model.

To test the mediation of the genetic and environmental influences by the latent personality factors, in the next step a common pathway and an independent pathway AE model were tested (Fig. 3). In the common pathway model, the variation in the latent five personality factors was decomposed into additive genetic and unique environmental components. Additive genetic influences explained around half of the variance in the latent traits (0.48, 0.48, 0.58, 0.43, and 0.47 for the N, E, O, A, and C factors, respectively), the remainder of the trait variance being determined by unique environmental factors. The fit measures associated with the model were: $\chi^2 = 112{,}786$, df = 14,776, TLI = 0.832, RMSEA = 0.069.[4] The independent pathway model was formulated by disposing of the phenotypic factors employed in the common pathway model. The fit measures associated with this model were: $\chi^2 = 94{,}852$, df = 14,721, TLI = 0.862, RMSEA = 0.062[4]. As the difference between Chi square values obtained using the WLSM estimator is not Chi square distributed (Muthén and Muthén 1998–2007), the comparison of the common and the independent pathway model was carried out using a Chi square difference test with scaling correction factors (Satorra and Bentler 2001). The resulting Chi square difference was $\Delta\chi^2 = 123{,}646$, df = 55. Additionally, the comparison was performed using maximum likelihood estimation with robust standard errors (MLR; Muthén and Muthén 1998–2007). The results converged with those obtained using the WLSM estimator (common pathway: $\chi^2 = 40{,}477$, df = 14,195, TLI = 0.699, RMSEA = 0.036; independent pathway: $\chi^2 = 35{,}423$, df = 14,140, TLI = 0.756, RMSEA = 0.033; Chi square difference: $\Delta\chi^2 = 3{,}115$, df = 55). The significant difference between the fit of the two models indicates incomplete mediation of the genetic and environmental influences by the latent personality factors.

In the light of the well-established presence of cross-loadings in the NEO-PI-R and the NEO-FFI (Parker et al. 1993; Block 1995; Costa and McCrae 2008), an additional test was performed: a common and an independent pathway model based on the phenotypic model with 94 cross-loadings were formulated and fitted to the data. Due to the computational intensity of fitting these models using the WLSMV estimator, the MLR estimator was used. The resulting fit measures were $\chi^2 = 31{,}176$, df = 14,101, TLI = 0.803, RMSEA = 0.029, and $\chi^2 = 25{,}831$, df = 13,952, TLI = 0.862, RMSEA = 0.025, respectively.

[4] As MPlus output obtained using the WLSMV estimator could not be used for subsequent Chi square difference testing due to the non-linear constraints in the model, estimation was performed using the WLSM estimator.

Table 4 Standardized factor loadings (λ), residual variances (μ_R = mean residual variance), and inter-factor correlations in the phenotypic 5-factor model

Item	λ_N	Res var	Item	λ_E	Res var	Item	λ_O	Res var
n1	0.52	0.73	e1	0.43	0.81	o1	0.32	0.90
n2	0.65	0.58	e2	0.61	0.63	o2	0.16	0.97
n3	0.63	0.60	e3	0.82	0.33	o3	0.61	0.63
n4	0.62	0.61	e4	0.58	0.66	o4	0.32	0.90
n5	0.73	0.47	e5	0.40	0.84	o5	0.70	0.51
n6	0.78	0.39	e6	0.33	0.89	o6	0.28	0.92
n7	0.63	0.61	e7	0.55	0.70	o7	0.35	0.88
n8	0.52	0.73	e8	0.85	0.28	o8	0.11	0.99
n9	0.68	0.54	e9	0.76	0.42	o9	0.76	0.42
n10	0.69	0.52	e10	0.08	0.99	o10	0.55	0.69
n11	0.70	0.51	e11	0.65	0.57	o11	0.47	0.78
n12	0.59	0.65	e12	0.33	0.89	o12	0.50	0.75
		$\mu_R = 0.58$			$\mu_R = 0.67$			$\mu_R = 0.78$
Item	λ_A	Res var	Item	λ_C	Res var			
e1	0.41	0.84	a1	0.45	0.80			
e2	0.47	0.78	a2	0.66	0.56			
e3	0.57	0.67	a3	0.40	0.84			
e4	0.36	0.87	a4	0.43	0.82			
e5	0.67	0.55	a5	0.51	0.74			
e6	0.48	0.77	a6	0.60	0.64			
e7	0.64	0.59	a7	0.58	0.67			
e8	0.42	0.83	a8	0.52	0.73			
e9	0.30	0.91	a9	0.54	0.71			
e10	0.55	0.70	a10	0.67	0.55			
e11	0.19	0.96	a11	0.84	0.30			
e12	0.31	0.91	a12	0.24	0.94			
		$\mu_R = 0.78$			$\mu_R = 0.69$			
Factor correlations:								
	N	E	O	A				
E	−0.62							
O	0.12	0.08						
A	−0.41	0.46	0.10					
C	−0.53	0.45	−0.03	0.48				

Consistently with the results obtained for the simple structure models, the fit of the two models differed significantly ($\Delta\chi^2 = 4{,}034$, df = 149), indicating incomplete mediation of the genetic and environmental effects by the latent personality factors, despite the assumption of simple structure being discarded.

Finally, to further explore the structures of the genetic and environmental influences on the item covariation, the 60 × 60 phenotypic polychoric covariance matrix was decomposed into 60 × 60 additive genetic and unique environmental matrices, and the dimensionality of these two covariance matrices was assessed using EFA (geomin rotation). The results are given in Fig. 4. As evident from the Figure, the scree plots (upper panel) for the A and the E matrix both indicate a 5-factor model. Furthermore, the factor structures of the additive genetic and the unique environmental influences (lower panel Fig. 4) resemble very closely the 5-factor phenotypic structure of the NEO-FFI. This can also be seen in Fig. 5, which depicts the pattern and the magnitude of the A and E intercorrelations between the items; as evident, the A and the E covariance structure resemble each other highly. Finally, the magnitudes of the A and E variance components of each of the 60 items are depicted in Fig. 6; on average, these are 0.33 and 0.67, respectively.

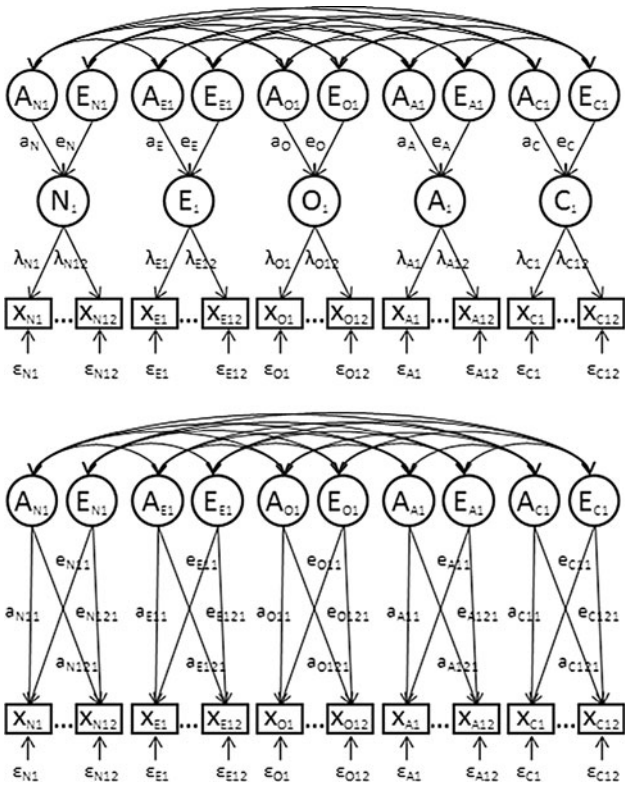

Fig. 3 The common (*upper panel*) and independent (*lower panel*) pathway models fitted to the NEO-FFI data. The models are only partially depicted; the full models include a 'twin1' and a 'twin2' part, analogous to Fig. 1. The 'within twin' A factors are mutually correlated, as are the 'within twin' E factors. The item-specific factors were modeled as correlated over twin 1 and twin 2 (i.e., the 60 × 60 twin 1–twin 2 residual covariance matrix is diagonal)

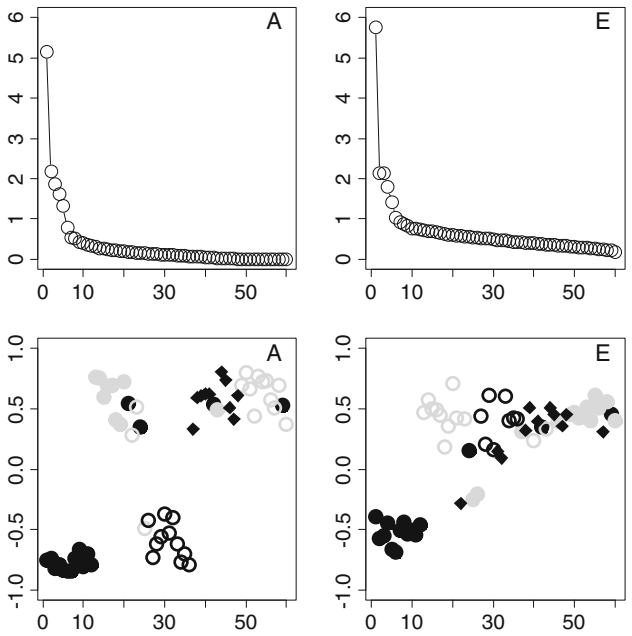

Fig. 4 Eigenvalues of $\mathbf{\Sigma}_A$ and $\mathbf{\Sigma}_E$ matrices (*upper panel*) and factor loadings obtained in EFA solutions with 5 A and 5 E factors (*lower panel*). *Shapes*/*shading* code for different latent factors. Only the highest factor loading for each item is shown

Discussion

In the present study, we tested the hypothesis that the Big FFs are causally efficient entities, which serve to mediate the genetic and environmental effects on the phenotypic data. This hypothesis was tested by comparing the fit of independent pathway models to the fit of common pathway models. If the latent variables in the FF model indeed act as causes of behavior, which fully mediate genetic effects, the independent and common pathway models should fit equally well. If, however, the latent variables are merely statistical constructs that organize phenotypic correlations but do not have the status of causally efficient entities, the independent pathway model should show superior fit. In addition to these hypothesis tests, the structures and the dimensionalities of the latent genetic and environmental effects were examined in an EFA. Two findings emerged: (1) the constraints associated with the common pathway model were not tenable, i.e., the fit indices favored the independent pathway model, and (2) the rotated 5-factor structures as obtained in the EFA of genetic and environmental correlation matrices are similar.

The fact that our analyses favor the independent pathway model constitutes evidence against the realist interpretation of the Big Five dimensions. Even when we allow cross-loadings to be present, the magnitude of the test statistic based on the models is such ($\Delta\chi^2 = 4{,}034$, df = 149, based on MLR) that the degree of misfit associated with the common pathway model is considerable. Perhaps, one could argue that both models fit well in view of the acceptable approximation error (common pathway model RMSEA: 0.029; independent pathway model RMSEA: 0.025). However, in our view, the acid test here is not the overall degree approximation error of the individual models. Rather, it is the model comparison, which reveals the specific source of approximation error, namely the proportionality constraints associated with the common pathway model. These are evidently untenable.

The fact that the exploratory factor analyses of the additive genetic and unshared environmental correlation matrices produced highly similar 5-factor models is interesting in its own right, and by no means a trivial finding. The phenotypic FF model does not imply five genetic and environmental factors to surface: the latter implies the former, but not vice versa, and several examples are known in which the structures diverge (Kendler et al. 1987; Franić et al. 2013b). Thus, although the data unambiguously reject the proportionality constraints derived from the latent variable hypothesis, it is certainly not the case that the A and E covariance structures are radically different.

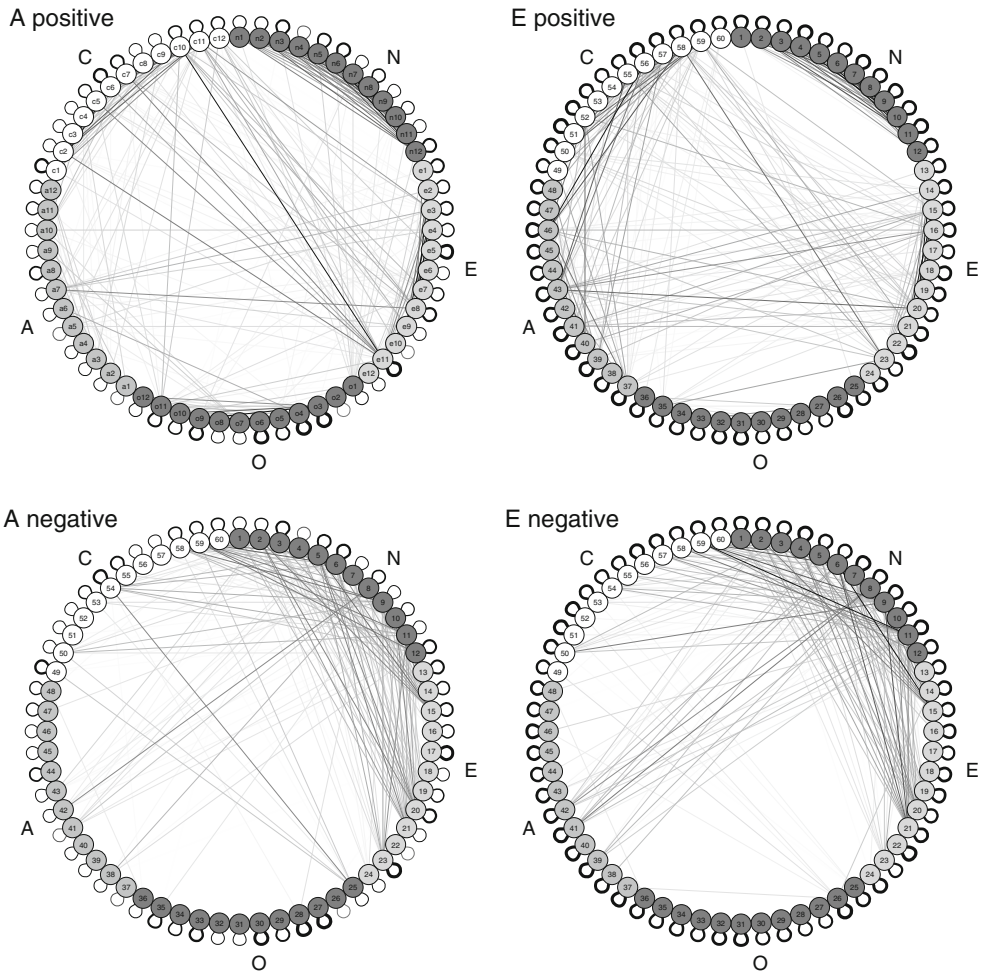

Fig. 5 Graphical representations (Epskamp et al. 2012) of the A (*left*) and E (*right*) covariance components of the NEO-FFI. Positive (*upper panel*) and negative (*lower panel*) covariances are shown separately. Nodes (i.e., *circles*) represent items. The thickness of the edges (i.e., of the *lines connecting the nodes*) represents the magnitude of covariance between the items

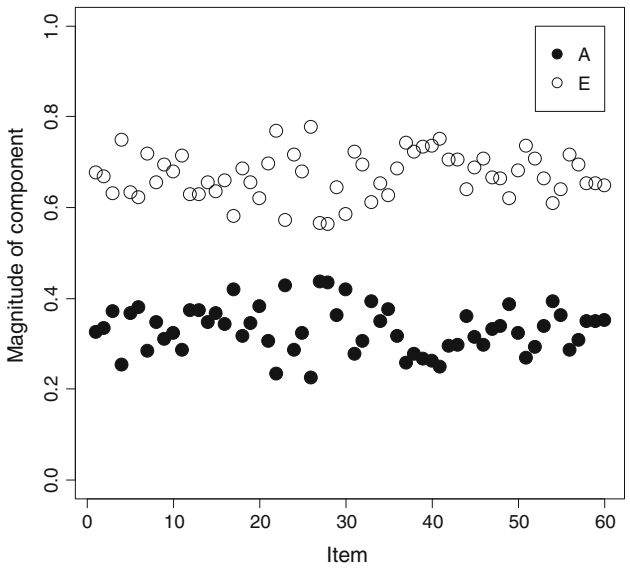

Fig. 6 Magnitude of the *A* and *E* variance components of the 60 items of NEO-FFI

Therefore, although the formal tests indicate that the independent pathway model is preferable, the exploratory results do lend some credence to the latent variable hypothesis. One possible explanation for this finding is that, although the full mediation hypothesis is not precisely true, it does provide a reasonable approximation to the generating model. The specific reasons for rejecting the common pathway model may, for instance, be highly local (due only to a subset of observed variables), and thus the violation may be accommodated by the addition of parameters or by the removal of offending variables. A second possible explanation is that, even though we have fitted highly relaxed versions of the FF model, the models still embodied auxiliary hypotheses that were not exactly true (e.g., linearity, normality, continuity, discarding C and D effects) which may have produced misfit evident in the likelihood ratio tests (which are derived on the hypothesis that the least restricted model fits the data). A third possibility is that the similar structure of A and E matrices, as evidenced in the present paper, is simply a chance finding

that has little to do with the realist interpretation of the Big Five dimensions. This hypothesis is tenable, because the truth of an independent pathway model does not preclude that the genetic and environmental covariance structures comprise 5 factors, with or without configurally similar loadings.

In our view, the formal test on the proportionality of loadings should carry the primary weight of the evidence, as it was designed specifically to distinguish between the tested models. However, it is certainly notable that the A and E covariance matrices showed strikingly similar structures, and even though this equivalence is not a formal test of the common pathway hypothesis, it does confirm an indirect consequence of that hypothesis. Further research may investigate the relevance of this finding to the veracity of the FF model.

In the present analyses, the genetic and environmental variables are all-encompassing in the sense that they represent all (unmeasured) polygenic and unshared environmental influences. However, the mediation hypothesis can be formulated with respect to any measured variable. It is a drawback of much of the research concerning the covariates of the Big Five dimensions that they generally involve Big Five subscale scores rather than items. We consider the demonstration of the mediatory role of, say, neuroticism in the relationship between a covariate (e.g., sex) and the neuroticism items, to be a stronger result than the demonstration of a sex difference in the neuroticism scale scores. In this regard the present results are relevant to gene-finding studies (e.g., genome-wide association studies; GWAS). If a measured genetic variant has its effect on the common factor "neuroticism", then its effect is present in all the relevant items, and the interpretation of the gene as a "gene for neuroticism" is tenable. This is not so if the effect is limited to a subset of the items, or perhaps even a single item.

References

Achenbach TM (1991) Manual for the child behavior checklist/4-18 and 1991 profile. Department of Psychiatry, University of Vermont, Burlington

Allport GW, Odbert HS (1936) Trait names: a psycho-lexical study. Psychol Monogr 47(1):1–211

Block J (1995) A contrarian view of the five-factor approach to personality description. Psychol Bull 117(2):187

Block JH, Block J (1980) The role of ego-control and ego-resiliency in the organization of behavior. In: Collins WA (ed) Minnesota symposium on child psychology, vol 13. Erlbaum, Hillsdale, pp 39–101

Bollen KA (1989) Structural equations with latent variables. Wiley, New York

Bouchard TJ Jr, Loehlin JC (2001) Genes, evolution, and personality. Behav Genet 31(3):243–273

Caprara GV, Barbaranelli C, Comrey AL (1995) Factor analysis of the Neo-PI inventory and the Comrey personality scales in an Italian sample. Personal Individ Differ 18(2):193–200

Cattell RB (1943a) The description of personality: basic traits resolved into clusters. J Abnorm Soc Psychol 38(4):476

Cattell RB (1943b) The description of personality. I. Foundations of trait measurement. Psychol Rev 50(6):559

Cattell RB (1945) The description of personality: principles [sic] findings in a factor analysis. Am J Psychol 58:69–90

Cervone D (2005) Personality architecture: within-person structures and processes. Annu Rev Psychol 56:423–452

Costa PT, McCrae RR (1985) The NEO personality inventory: manual, form S and form R. Psychological assessment resources, Odessa

Costa PT, McCrae RR (1992) Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Psychological Assessment Resources Inc., Odessa

Costa P, McCrae R (2008) The revised NEO personality inventory (NEO-PI-R). The SAGE handbook of personality theory and assessment, pp 2179–2198

de Moor MH, Costa P, Terracciano A, Krueger R, De Geus E, Toshiko T, Penninx B, Esko T, Madden P, Derringer J (2010) Meta-analysis of genome-wide association studies for personality. Mol Psychiatry 17(3):337–349

DeYoung CG, Hirsh JB, Shane MS, Papademetris X, Rajeevan N, Gray JR (2010) Testing predictions from personality neuroscience brain structure and the big five. Psychol Sci 21(6):820–828

Dolan CV (1994) Factor analysis of variables with 2, 3, 5 and 7 response categories: a comparison of categorical variable estimators using simulated data. Br J Math Stat Psychol 47(2):309–326

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D (2012) qgraph: network visualizations of relationships in psychometric data. J Stat Softw 48(4):1–18

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longmans Green, Harlow

Flora DB, Curran PJ (2004) An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods 9(4):466–491

Franić S, Dolan CV, Borsboom D, Boomsma DI (2012) Structural equation modeling in genetics. In: Hoyle RH (ed) Handbook of structural equation modeling. Guilford Press, New York, pp 617–635

Franić S, Dolan CV, Borsboom D, Hudziak JJ, van Beijsterveldt CEM, Boomsma DI (2013a) Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. Psychol Methods 18(3):406–433

Franić S, Dolan CV, Borsboom D, van Beijsterveldt CEM, Boomsma DI (2013b) Three-and-a-half-factor model? The genetic and environmental structure of the CBCL/6-18 internalizing grouping. doi:10.1007/s10519-013-9628-4

French JW (1953) The description of personality measurements in terms of rotated factors. Educational Testing Service, Princeton, NJ

Goldberg LR (1977) Language and personality: developing a taxonomy of trait descriptive terms. Invited address to the division of evaluation and measurement at the annual meeting of the American psychological association, San Francisco

Goldberg LR (1980) Some ruminations about the structure of individual differences: developing a common lexicon for the major characteristics of personality. Paper presented at the annual meeting of the western psychological association, Honolulu, HI

Goldberg LR (1981) Language and individual differences: the search for universals in personality lexicons. Rev Personal Soc Psychol 2(1):141–165

Goldberg LR (1982) From ace to zombie: some explorations in the language of personality. Adv Personal Assess 1:203–234

Goldberg LR (1983) The magical number five, plus or minus two: Some conjectures on the dimensionality of personality descriptions. Paper presented at a research seminar, Gerontology Research Center, Baltimore, MD
Goldberg LR (1990) An alternative description of personality—the big-5 factor structure. J Personal Soc Psychol 59(6):1216–1229
Goldberg LR (1992) The development of markers for the big-five factor structure. Psychol Assess 4(1):26–42
Goldberg LR (1993) The structure of phenotypic personality traits. Am Psychol 48:26–34
Hahn R, Comrey AL (1994) Factor analysis of the NEO-PI and the Comrey personality scales. Psychol Rep 75(1):355–365
Hoekstra HA, Ormel J, De Fruyt F (1996) NEO personality questionnaires NEO-PI-R, NEO-FFI: manual. Swet & Zeitlinger BV, Lisse
Johnson W, Krueger RF (2004) Genetic and environmental structure of adjectives describing the domains of the big five model of personality: a nationwide US twin study. J Res Personal 38(5):448–472
Jöreskog KG (1993) Testing structural equation models. In: Bollen KA, Long SJ (eds) Testing structural equation models. SAGE, Newbury Park, pp 294–316
Jöreskog KG, Sörbom D (2004) LISREL. Scientific Software International, Inc., Skokie
Keller MC, Coventry WL (2005) Quantifying and addressing parameter indeterminacy in the classical twin design. Twin Res Hum Genet 8(3):201–213
Keller MC, Medland SE, Duncan LE (2010) Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. Behav Genet 40(3):377–393
Kendler KS, Heath AC, Martin NG, Eaves LJ (1987) Symptoms of anxiety and symptoms of depression—same genes, different environments. Arch Gen Psychiatry 44(5):451–457
Kenny DA (2012) Measuring model fit. http://davidakenny.net/cm/fit.htm. Accessed 28 June 2013
Kline RB (2005) Principles and practice of structural equation modeling. Guilford Press, New York
Loehlin JC (1989) Partitioning environmental and genetic contributions to behavioral development. Am Psychol 44(10):1285
Loehlin J, Martin N (2001) Age changes in personality traits and their heritabilities during the adult years: evidence from Australian twin registry samples. Personal Individ Differ 30(7):1147–1160
Markus K, Borsboom D (2013) Frontiers of validity theory: measurement, causation, and meaning. Routledge, New York
Martin NG, Eaves LJ (1977) Genetic-analysis of covariance structure. Heredity 38:79–95
Mather K, Jinks JL (1971) Biometrical genetics. Chapman and Hall, London
McArdle JJ, Goldsmith HH (1990) Alternative common factor models for multivariate biometric analyses. Behav Genet 20(5):569–608
McCrae RR, Costa PT Jr (1983) Joint factors in self-reports and ratings: neuroticism, extraversion and openness to experience. Personal Individ Differ 4(3):245–255
McCrae RR, Costa PT (1999) A five-factor theory of personality. Guilford, New York
McCrae RR, Costa PT (2008) Empirical and theoretical status of the five-factor model of personality traits. Sage handbook of personality theory and assessment, vol 1. Sage Publications, Thousand Oaks, CA, pp 273–294
McCrae RR, John OP (1992) An introduction to the 5-factor model and its applications. J Personal 60(2):175–215
McCrae RR, Zonderman AB, Costa PT, Bond MH, Paurnonen S (1996) Evaluating replicebility of factors in the revised NEO personality inventory: confirmatory factor analysis versus procrustes rotation. J Personal Soc Psychol 70:552–566
Mellenbergh GJ (1989) Item bias and item response theory. Int J Educ Res 13(2):127–143
Meredith W (1993) Measurement invariance, factor-analysis and factorial invariance. Psychometrika 58(4):525–543
Mroczek DK (1992) Personality and psychopathology in older men: the five factor model and the MMPI-2. Dissert Abstr int 53(4B):2095
Muthén LK, Muthén BO (1998–2007) Mplus user's guide. Muthén & Muthén, Los Angeles
Neale MC (2000) MxGui (1.7.03) [Computer software]. Virginia Commonwealth University, Richmond
Neale MC, Cardon L (1992) Methodology for genetic studies of twins and families. Kluwer Academic Publishers B.V, Dordrecht
Nigg JT, John OP, Blaskey LG, Huang-Pollock CL, Willicut EG, Hinshaw SP, Pennington B (2002) Big five dimensions and ADHD symptoms: links between personality traits and clinical symptoms. J Personal Soc Psychol 83(2):451
Norman WT (1963) Toward an adequate taxonomy of personality attributes: replicated factor structure in peer nomination personality ratings. J Abnorm Soc Psychol 66(6):574
Norman WT (1967) 2800 personality trait descriptors: normative operating characteristics for a university population. Department of Psychological Sciences, University of Michigan, Ann Arbor
Parker JD, Bagby RM, Summerfeldt LJ (1993) Confirmatory factor analysis of the revised NEO personality inventory. Personal Individ Differ 15(4):463–466
Peabody D, Goldberg LR (1989) Some determinants of factor structures from personality-trait descriptors. J Personal Soc Psychol 57(3):552–567
Pervin LA (1994) Further reflections on current trait theory. Psychol Inq 5(2):169–178
Plomin R, Caspi A (1990) Behavioral genetics and personality. Handbook of personality: theory and research, vol 2. Guilford Press, New York, pp 251–276
R Development Core Team (2009) R: a language and environment for statistical computing. In: R Foundation for Statistical Computing, Vienna, Austria
Rebollo I, de Moor MHM, Dolan CV, Boomsma DI (2006) Phenotypic factor analysis of family data: correction of the bias due to dependency. Twin Res Hum Genet 9(3):367–376
Satorra A, Bentler PM (2001) A scaled difference Chi square test statistic for moment structure analysis. Psychometrika 66(4):507–514
Schmit MJ, Ryan AM (1993) The Big five in personnel selection: factor structure in applicant and nonapplicant populations. J Appl Psychol 78(6):966
Tupes EC, Christal RE (1992) Recurrent personality factors based on trait ratings. J Personal 60(2):225–251
van Dongen JP, Draisma HHM, Martin NG, Boomsma DI (2012) The continuing value of twin studies in the omics era. Nat Rev Genet 13(9):640–653
Willemsen G, Vink JM, Abdellaoui A, den Braber A, van Beek JHDA, Draisma HHM, van Dongen J, van't Ent D, Geels LM, van Lien R (2013) The adult Netherlands twin register: twenty-five years of survey and biological data collection. Twin Res Hum Genet 16(1):271–281