Against All Odds: Bifactors in EFAs of Big Five Data

Michael Biderman

University of Tennessee at Chattanooga

Part of symposium: S. McAbee & M. Biderman, Chairs. Theoretical and Practical Advances in Latent Variable Models of Personality. Conducted at the 29th annual conference of The Society for Industrial and Organizational Psychology; Honolulu, Hawaii, 2014.

Authors' Note: Correspondence regarding this article should be sent to Michael Biderman, Department of Psychology / 2803, U.T. Chattanooga, 615 McCallie Ave., Chattanooga, TN 37403. Tel.: (423) 316-9504. Email: Michael-Biderman@utc.edu.

This paper and others related to it may be downloaded from www.utc.edu/Michael-Biderman

Against All Odds: Bifactors in EFAs of Big Five Data

Michael Biderman

University of Tennessee at Chattanooga

The dimensions of personality measured by Big Five questionnaires – extraversion, agreeableness, conscientiousness, neuroticism/stability, and openness/intellect - have become a cornerstone of research since first being popularized in the 1990s (Costa & McCrae, 1992; Digman, 1990; Goldberg, 1993). The dimensions have been found in nearly all cultures worldwide. Very high convergent validity has been found across multiple Big Five questionnaires. One of the most popular of those questionnaires is the NEO-FFI (Costa & McCrae, 1992). It is the focus of this presentation.

A large majority of analyses involving the Big Five have used summated scales computed from the item responses. The NEO-FFI has 60 items, split into five groups of 12 items each. Thus each NEO-FFI Big Five scale score is based on the sum of 12 items. Although analyses involving such scales have yielded much useful information about personality, it is possible that summating across items may obscure important information about behavior contained in the individual items. There may be personality characteristics affecting individual item responses whose effects are masked when the items are combined into scale scores. The focus of the present paper is on modeling of individual item data.

Figure 1 presents a Confirmatory Factor Analysis (CFA) of the 60 items of the NEO-FFI. The figure shows that each item is influenced by only one personality characteristics –the specific Big Five factor on which the item loads. Symbols representing residual effects – errors of measurement – have been omitted from the figure. Interfactor correlations are included in the figure based on evidence from many studies of the Big Five showing small correlations between the factors. When the neuroticism/stability factor is scored as stability, generally all the correlations are positive, with mean about 0.1. Correlations between items indicating different factors are represented in CFA models such as that of Figure 1 by paths that include correlations between factors.

Figure 2 presents a bifactor CFA model of the NEO-FFI items. The right side of this figure is the same model as shown in Figure 1. On the left side of Figure 2, however, a sixth first order factor has been added. All items load on the sixth factor. Such a factor on which all items load, so that each item is influenced by two factors, is called the general factor in bifactor models (Holzinger & Swineford, 1937). In such models, the other first order factors on which just some items load are called group factors. The general factor is labeled M in Figure 2 because of its resemblance in the model to factors called method factors. Since it influences all items of the questionnaire, if the loadings of all items on it were positive, individual differences in M would be expected to lead to positive correlations among all items in the questionnaire. These positive correlations would be reflected in common variance of all the questionnaire items.

The need for a factor that could account for common variance across items was first suggested in studies of faking of Big Five questionnaires. In those studies, we found that instructions or incentives to fake typically resulted in level shifts in responses to all items of a questionnaire, level shifts that seemed at the time (and still do) to be best accounted for by assuming that there was a separate “faking” tendency, i.e., factor, affecting responses in faking conditions. This belief was supported by the result that adding this factor as an influence on faked items resulted in significant improvements on goodness-of-fit (Biderman & Nguyen, 2004; Wrensen & Biderman, 2005.)

Our original belief was the there was no need for such a factor in models of data in which persons were instructed to respond honestly - that there was little common variance in honest item responses, so little that adding such a factor would be unproductive. Curiosity ruled, however, and we investigated bifactor models of data in which people were instructed to respond honestly. We were surprised to find when analyzing the data of the 50-item questionnaire provided free-of-charge on the IPIP web site (ipip.uni.org; Goldberg, 1999) that adding a general factor to the model of honest response data resulted in significant increases in goodness-of-fit of the model just as it did to models of data gathered under faking instructions or incentives (Biderman, 2007).

Our study of the importance of bifactor CFA models of Big Five data obtained under instructions to respond honestly was summarized by Biderman, Nguyen, Cunningham, & Ghorbani (2011). In that study adding a general factor significantly improved goodness-of-fit in five separate data sets in which the IPIP 50-item questionnaire had been administered. More importantly for the present study, we also found in a sixth data set that adding a general factor to a CFA model of NEO-FFI data resulted in a significant improvement in goodness-of-fit to those data also. This presentation follows up on the analyses conducted by Biderman et al. (2011). It uses the NEO-FFI data analyzed by them. The details of the data set are available in that paper. The results of the comparison of goodness-of-fit measures for the NEO-FFI CFA in Biderman et al. (2011) are presented in Table 1. From these results we conclude that adding a general factor to a correlated factors CFA of the NEO-FFI data yielded significantly improved goodness-of-fit.

Although the confirmation of the need for a general factor in CFA models in analyses across data sets and across questionnaires was comforting, as is the case in any discovery such as this, alternative explanations of the same result must be considered. Two such alternatives are considered here, and are the primary focus of this presentation. The first is the possibility that in spite of the efforts of developers of the Big Five questionnaires to select items that represent only one Big Five factor – the factor on which the items load in Figure 1 – in fact, some, perhaps all of the items, are influenced by multiple Big Five factors. That is, perhaps cross-loadings of items onto all of the Big Five factors should be allowed. Those cross-loadings would certainly account for much common variance in the items, perhaps so much so that there would be no need for a sixth factor on which all items load – the general factor. A model in which all cross loadings are estimated is what is commonly known as an exploratory factor analysis or EFA model. An EFA model of the NEO-FFI items without a general factor is presented in Figure 3. All factors are allowed to correlate. The same model with a general factor added is presented in Figure 4.

The models were applied using the ESEM feature recently added to Mplus (Asparouhov & Muthén, 2009). This feature allow the estimation of EFA models while at the same time permitting estimation of model parameters that are typically only available in software designed to apply CFA or SEM models. The ESEM option results in estimates that are virtually identical to those of an EFA model if no other model parameters are estimated. The program also contains a rotation option that allows one factor in an EFA to be treated as a general factor such as shown in Figure 4. The results of the comparison of the model of Figure 3 with that of Figure 4 are presented in Table 2. As the table shows, the bifactor EFA model fit the data significantly better than did the simple oblique factors EFA model. Thus, adding a sixth general factor to an EFA model of NEO-FFI data results in significant improvement in goodness-of-fit.

In order to insure that the general factor estimated from the EFA model was the same factor as that estimated from the CFA model, the correlation of standardized loadings of the items on the EFA and the CFA general factors was computed. That correlation was .966. A plot of the EFA general factor loadings vs. the CFA general factor loadings is presented in Figure 5.

The second challenge to the need for a bifactor model involves the possibility that common variance in the items could be the result of idiosyncratic correlations between pairs of items. Such correlations are typically represented by adding correlated residuals to CFA models. For a 60-item questionnaire, there are 1770 possible correlated residuals that could be estimated. Clearly some rationale for estimating a subset of that number must be provided. Such a rationale was presented by Marsh, Lüdtke, Muthén, Asparouhov, Morin, & Trautwein (2010). They identified 57 pairs of items in the NEO-FFI that represent identical facets in the NEO-PI-R questionnaire, from which the NEO-FFI items were taken. Marsh et al. (2010) found that including those 57 correlated residuals resulted in substantial improvement in goodness-of-fit of an oblique factors model of NEO-FFI data. Residual correlations between the same 57 pairs of items presented in the Appendix to Marsh et al. (2010) were added to the EFA models being considered here. Figure 6 presents the model without a general factor. Figure 7 presents the model with a general factor. It should be noted that the correlated residuals symbols in the two figures are not the actual pairs correlated but were chosen to symbolize the actual 57 pairs. Since the pairs that were correlated represented identical facets, none of the correlated residuals involved items indicating different Big Five factors.

The Mplus ESEM procedure was used to estimate the models of Figures 6 and 7. The results are presented in Table 3. As shown in Table 3 adding a general factor to the model resulted in a significant improvement in goodness-of-fit. To insure that the general factor estimated in this more complex model was the same factor as that estimated in the original CFA model, the correlation of standardized loadings of items on the general factor between the original CFA model and the correlated residuals EFA model was computed. It was .955. Figure 8 presents a plot of the EFA+Correlated Residuals loadings vs. the original CFA loadings.

To summarize, the results above show that adding a general factor on which all items load – creating a bifactor model - significantly improves goodness-of-fit in traditional CFA models, in EFA models, and in EFA models with rationally chosen correlated residuals estimated. A natural follow-up question concerns the importance of these results in modeling Big Five data. We offer two reasons for researchers to conduct future research on bifactor models of Big Five data.

The first reason that bifactor models should be considered further is that the general factor estimated here may represent a stable characteristic of personality, one on a par with the other Big Five. Our research has suggested that this characteristic might be what would be called a general expression of affective state. That is, persons high on the general factor are responding in a fashion to indicate that they have a generally positive view of themselves. Persons responding lower on the general factor may be responding to indicate that they have a generally negative view of themselves. Evidence for this view comes from correlations of the general factor with measures of affect. Biderman et al. (2011; Table 7) presented evidence that the general factor estimated from the IPIP 50-item questionnaire was positively correlated with measures of positive affect and negatively correlated with measures of negative affect. We have just completed gathering comparable data from the NEO-FFI-3, an updated version of the NEO-FFI questionnaire on which the above results are based. As shown in Table 4, we have found the following correlations of the general factor estimated from a bifactor model of the NEO-FFI-3: Costello and Comrey Depression scale: -.486; PANAS Negative Affectivity scale: -.249; Rosenberg Self-esteem scale: +.461, and PANAS Positive Affectivity scale: +.622. Sample sizes varied from 111 to 352. All were significantly different from 0. Thus, these results for the NEO-FFI-3 replicate results from the IPIP questionnaire reported by Biderman et al. (2011).

The second reason for considering the general factor is that regardless of what it is, it seems to exist and is thus a contaminant of Big Five scale scores. For that reason, if for no other reason, it should be partialled out of relationships involving the Big Five scale scores. Table 5 presents correlations of NEO-FFI-3 scale scores with the Costello and Comrey depression scale scores, first without partialling the general factor and then after partialling the general factor. The unpartialled correlations of all Big Five scale scores except openness were significantly negative. After partialling, all negative correlations were less negative, and two of the four were not significant. This pattern of results replicates results found by Biderman et al. (2011) for the IPIP questionnaire. Thus, even if there is disagreement on what the bifactor estimated from Big Five data represents, it clearly affects scale scores computed from Big Five data and for that reason, it would seem prudent to estimate it and to partial it from relationships involving those scores.

In conclusion, this paper has presented evidence that responses to a common Big Five questionnaire – the NEO-FFI – are influenced by a general factor. The evidence presented here suggests the general factor should be considered in CFA models, in EFA models, and in EFA models with correlated residuals. The factor may represent a stable personality characteristic. The evidence suggests that correlations involving Big Five measures should partial out its effects.

References.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397-438.

Biderman, M. D., & Nguyen, N. T. (2004). Structural equation models of faking ability in repeated measures designs. Paper presented at the 19th Annual Society for Industrial and Organizational Psychology Conference, Chicago, IL.

Biderman, M. D. (2007). Method variance and big-five correlations. Paper presented at the 7th Annual Conference of the Association for Research in Personality, Memphis, TN, 2007, January 24, 2007.

Biderman, M. D., Nguyen, N. T., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in Personality, 45*, 417-429.

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

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

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist 48*, 26–34.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F.

Ostendorf (Eds.), *Personality Psychology in Europe, Vol. 7* (pp. 1-28). Tilburg, The Netherlands: Tilburg University Press.

Holzinger, K. J., & Swineford, R. (1937). The befactor method. *Psychometrika, 2*, 41-54.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the Big Five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471-491.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Wrensen, L. B., & Biderman, M. D. (2005). Factors related to faking ability: A structural equation model application. Paper presented at the 20th Annual Conference of The Society for Industrial and Organizational Psychology, Los Angeles, CA. April.

Table 1. Comparing the CFA models from Biderman et al. (2011). N=196.

Model	Chi-square	df	CFI	TLI	RMSEA	SRMR
CFA	3223.910	1700	0.639	0.624	0.068	0.093
Bifactor CFA	2897.786	1640	0.702	0.679	0.063	0.075
Difference	326.124	60	0.063	0.055	0.005	0.018

Table 2. Comparing EFA models. N=196.

Model	Chi-square	df	CFI	TLI	RMSEA	SRMR
EFA	2494.515	1480	0.760	0.713	0.059	0.051
Bifactor EFA	2327.143	1425	0.786	0.735	0.057	0.047
Difference	167.372	55	0.026	0.022	0.002	0.004

Table 3. Comparing EFA models in which Marsh et al. (2010) EFA correlated Residuals are included in the models. N=196.

Model	Chi-square	df	CFI	TLI	RMSEA	SRMR
Corr’d Res	2084.048	1423	0.843	0.805	0.049	0.047
Corr’d Res Bifactor	1940.077	1368	0.865	0.825	0.046	0.044
Difference	143.971	55	0.022	0.020	0.003	0.003

Table 4. Correlations of bifactor estimated from CFA+bifactor model applied to NEO-FFI-3 data with measures of affect.

Depression	NA	RSE	PA
-.486	-.249	+.461	+.622

Table 5. Effects of partialling out the bifactor on correlations of NEO-FFI-3 scale scores with measures of affect. Correlations in red are significantly different from zero (p < .05).

Unadjusted correlations of Big Five Scales with Depression

Ext	Agr	Con	Sta	Opn
-.581	-.290	-.219	-.441	-.063

Correlations after partialling the bifactor

Ext	Agr	Con	Sta	Opn
-.404	-.165	-.005	-.420	.097

Figure 1. A CFA model of NEO-FFI items. Neuroticism has been reversed-scored and is labeled Stability.

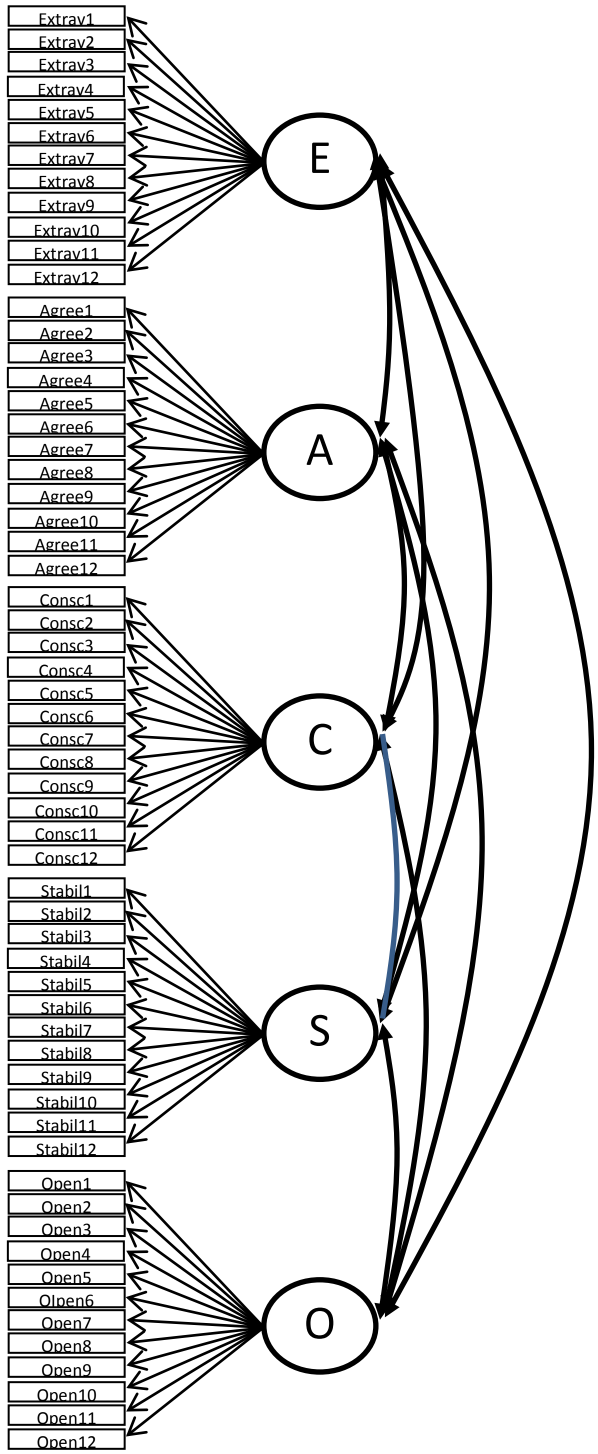

Figure 2 . A bifactor CFA model of NEO-FFI items.

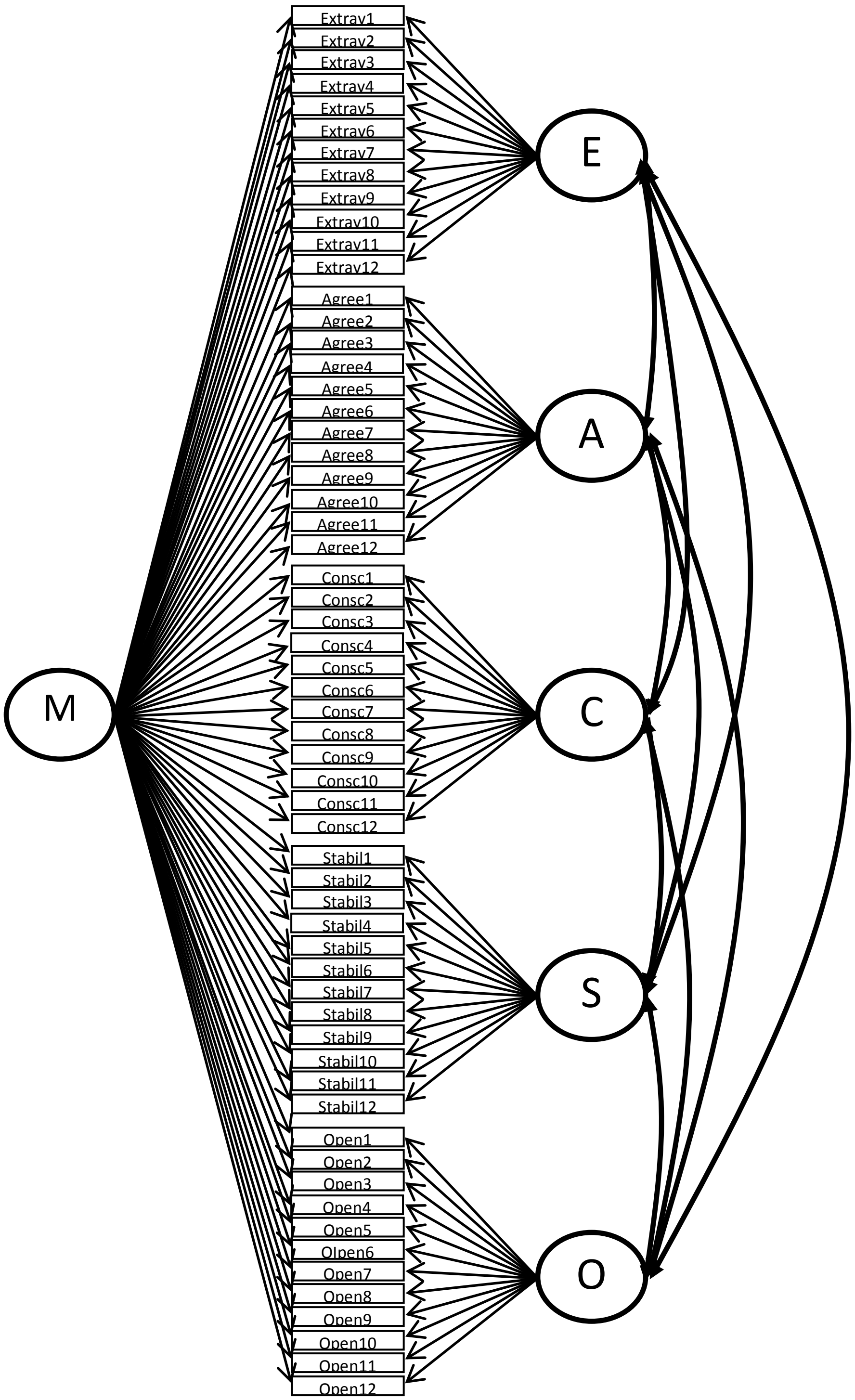

Figure 3. An oblique factors EFA model of NEO-FFI items.

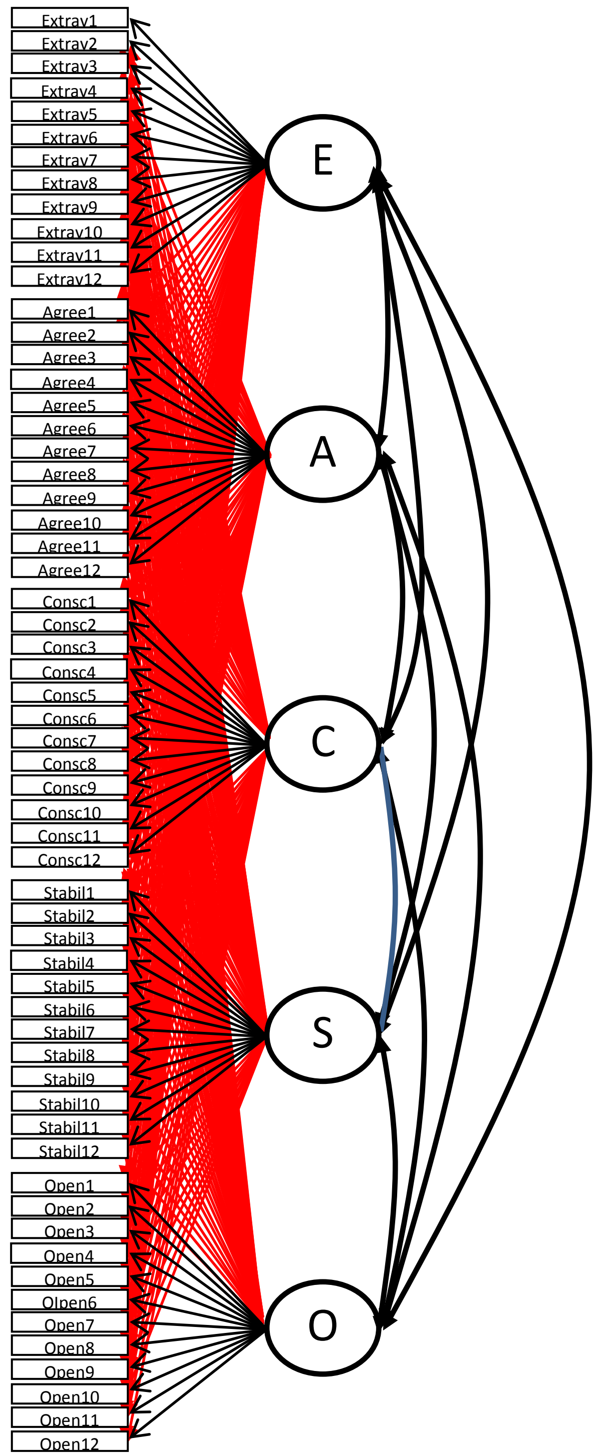

Figure 4. A bifactor oblique factors EFA model of NEO-FFI items.

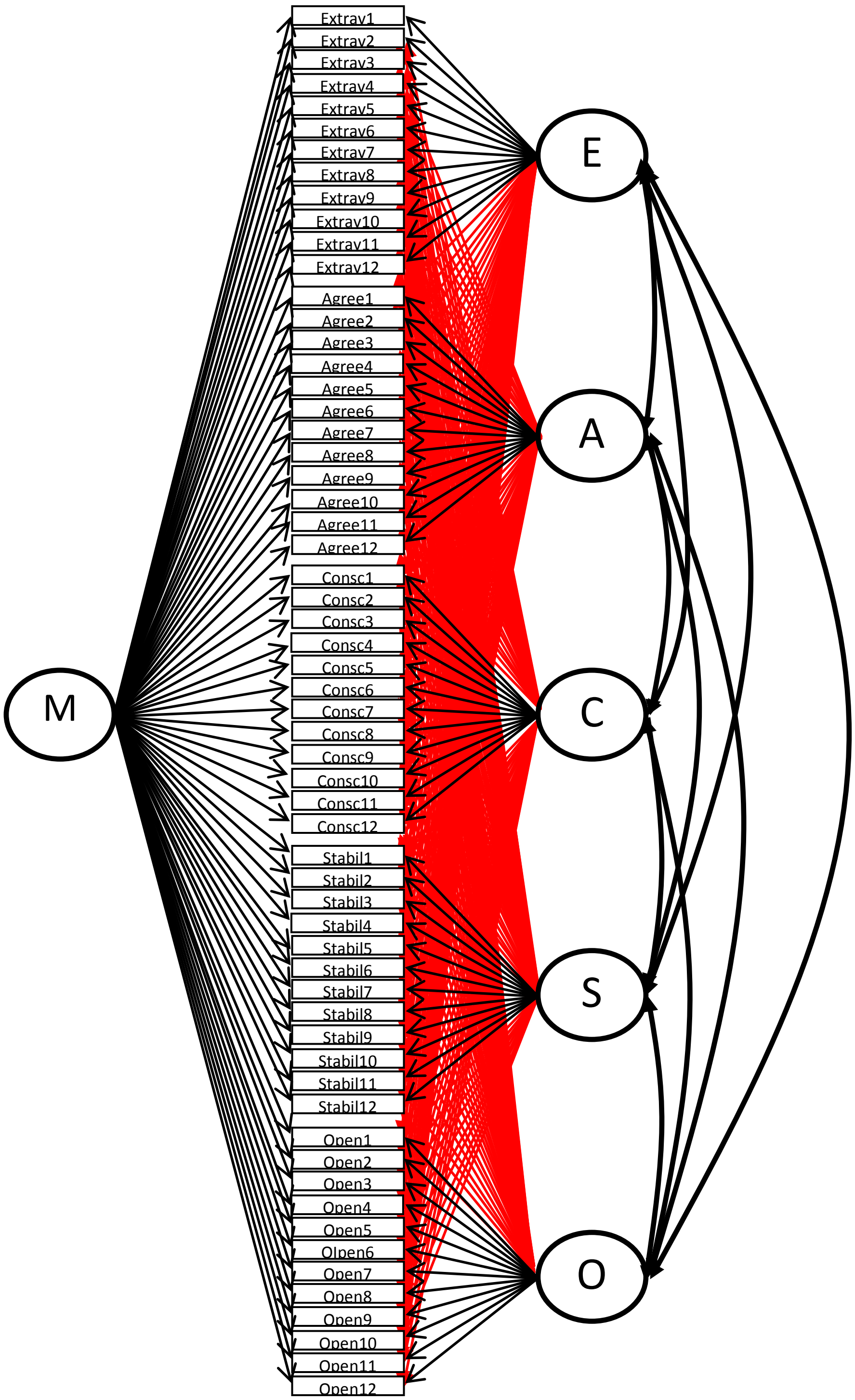

Figure 5. Standardized loadings of NEO-FFI items on the general factor from the bifactor EFA model vs. standardized loadings of the items from the bifactor CFA model.

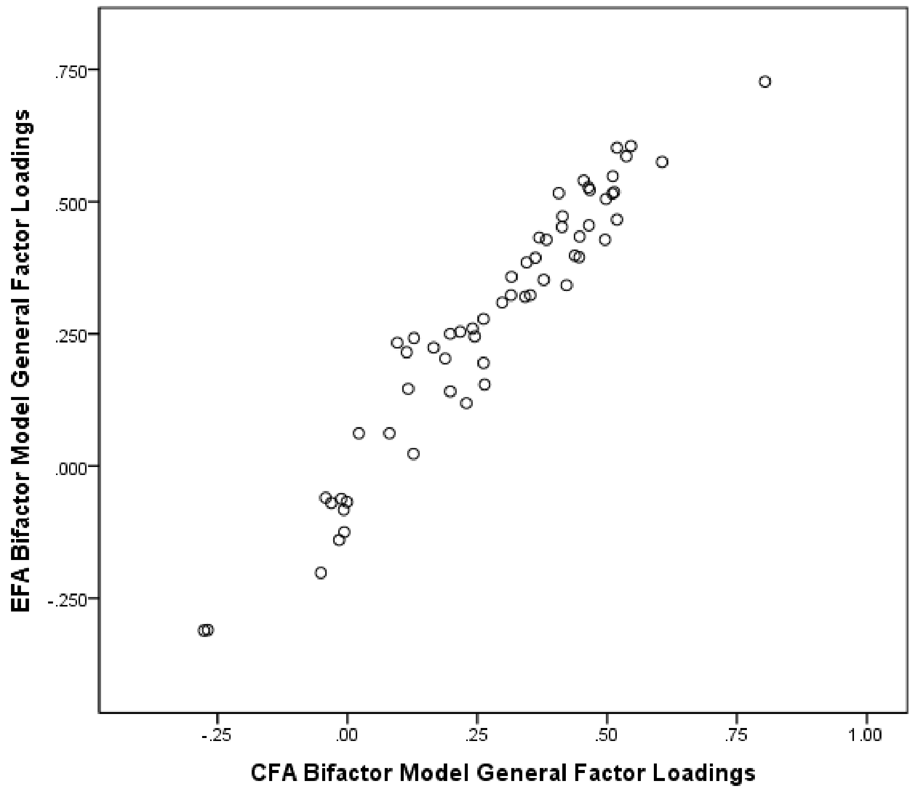

Figure 6 An oblique factors EFA model of the NEO-FFI with correlated residuals. The arrows representing correlated residuals are only suggestive - the items shown in the figure connected by arrows are not the actual items whose residuals were correlated.

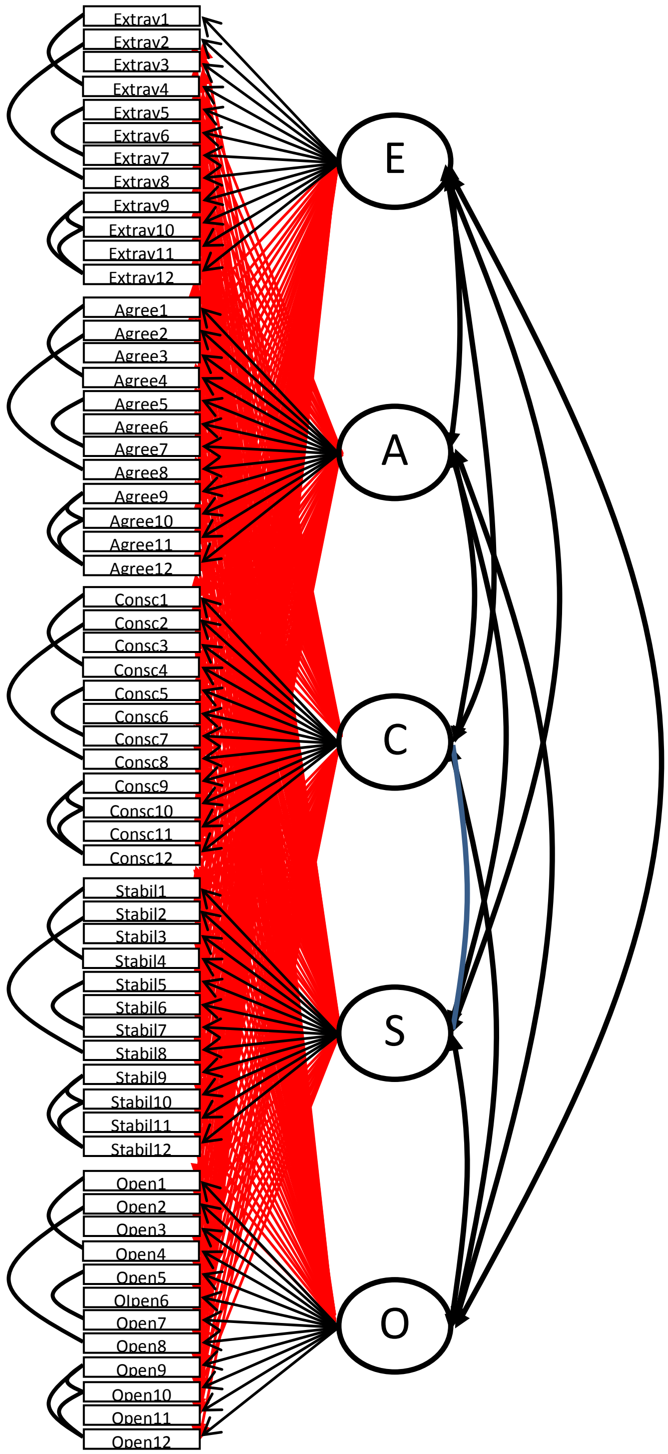

Figure 7 An oblique factors bifactor EFA model of the NEO-FFI with correlated residuals. The arrows representing correlated residuals are only suggestive - the items shown in the figure connected by arrows are not the actual items whose residuals were correlated.

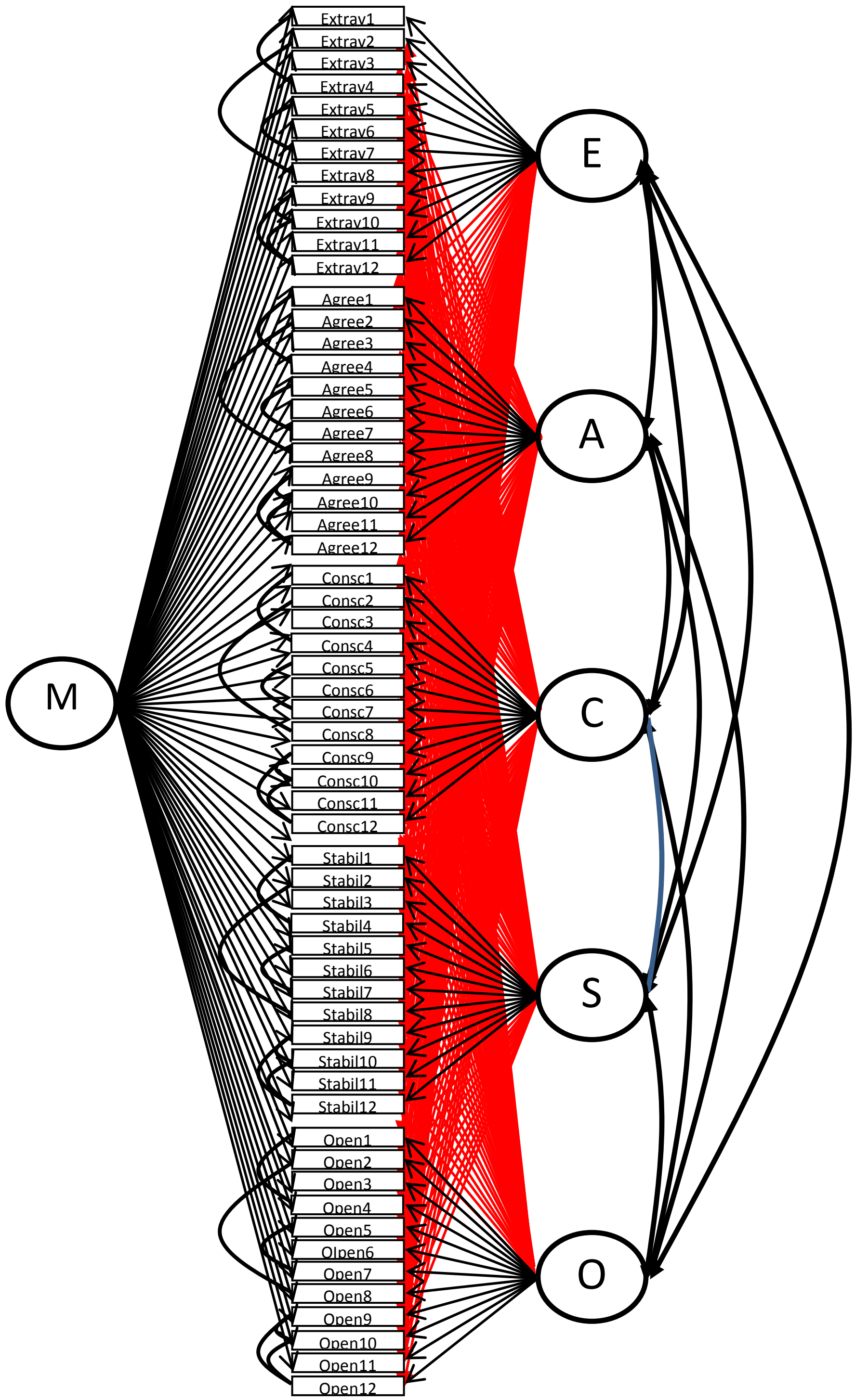

Figure 8. Standardized loadings of NEO-FFI items on the general factor from the bifactor EFA Correlated Residuals model vs. standardized loadings of the items from the bifactor CFA model.