Personality and Individual Differences 117 (2017) 122–129

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Psychometric properties and comparison of different techniques for factor analysis on the Big Five Inventory from a Flemish sample

Anikó Lovik [a,*], Vahid Nassiri [a], Geert Verbeke [a,b], Geert Molenberghs [b,a], An Katrien Sodermans [c]

[a] I-BioStat, KU Leuven - University of Leuven, Leuven, Belgium
[b] I-BioStat, Hasselt University, Diepenbeek, Belgium
[c] Centre for Sociological Research, University of Leuven, Leuven, Belgium

ARTICLE INFO

Article history:
Received 29 March 2017
Received in revised form 29 May 2017
Accepted 30 May 2017
Available online 3 June 2017

Keywords:
Big Five Inventory
Divorce in Flanders
Factor analysis
Factor rotation
Ipsatisation
Listwise deletion
Missing data
Multiple imputation
Ordinal variables
Polychoric correlation

ABSTRACT

In this paper we examine the Dutch language version of the Big Five Inventory, a short questionnaire used to measure the Big Five personality factors, on a Flemish sample coming from the Divorce in Flanders study. Our aim is twofold. First, we show that based on the Flemish sample the Dutch BFI has good psychometric properties and a clear factor structure comparable to a previous Dutch sample and in the international Big Five research literature.

Second, we compare the usual method of analysis, namely factor analysis with principal component extraction with varimax rotation, to several methods that each address a common problem in factor analysis. We compare the original analysis to factor analysis with a non-orthogonal rotation (addressing the problem of correlated factors), after ipsatisation (considering individual response styles), using polychoric correlations (taking into account the type of the responses), and using multiple imputation to handle missingness (to account for potential bias due to listwise deletion). The five factor analyses do not differ substantially. However, the analysis using polychoric correlations has the highest factor loadings and explains more of the variance than any other analysis; the analysis using ipsatised scores provides the worst results in supporting the Big Five structure.

1. Introduction

Questionnaires have been favoured tools in psychological and sociological research from the beginning of the last century for several reasons. Among others, they allow for relatively cheap and fast data collection and the possibility to reach populations that are unavailable via other methods. The increasing popularity of survey methods resulted in the development of statistical methods that allow reliable ways for analysing the collected data. Factor analysis has been a widely used method in social sciences to handle the massive amounts of survey data since the early days – where typically researchers want to explore the underlying, unobservable structure beyond the items assessed in the questionnaires. From the first half of the 20th century more and more personality questionnaires were developed to help clinicians and researchers to describe and understand the structure of personality (McCrae & John, 1992). The emergence of the Big Five model of personality (Digman, 1990; McCrae & Costa, 1987) resulted in the development of several questionnaires (for a review see Widiger & Trull, 1997; John & Srivastava, 1999) such as the Big Five Inventory (John & Srivastava, 1999).

While there are several other personality models in use, the Big Five is the most established and best validated model of personality and can be measured with personality tests where each of the five factors is assessed with a set of items (John & Srivastava, 1999; McCrae & Costa, 1987). The five factors are Neuroticism (N), Extraversion (E), Openness to Experience (O), Conscientiousness (C), and Agreeableness (A) (Costa & McCrae, 1992; Digman, 1990; John & Srivastava, 1999; McCrae & Costa, 1987). Neuroticism is characterised by anxiety, nervosity, sadness and is the polar opposite of emotional stability. Extraversion is linked to sociability, assertiveness, and energy. Openness to Experience refers to originality, curiosity, creativity and intelligence. Conscientiousness is related to orderliness, responsibility, and dependability. Agreeableness implies characteristics such as good-naturedness, modesty, compliance, cooperativeness, and trust (John & Srivastava, 1999).

The five factors are generally found across cultures (Hofstee, Kiers, de Raad, Goldberg, & Ostendorf, 1997) and do not change considerably with age (Asendorpf & van Aken, 2003). Thus, to acquire Big Five questionnaires, translations of existing questionnaires are used and validated by comparing the results to other translations or similar questionnaires.

Although there are several well-known questionnaires to measure the Big Five factors, most of them are rather long, which can seriously limit their usability as most survey research has to take into account

* Corresponding author at: L-BioStat, Kapucijnenvoer 33 blok d, 3000 Leuven, Belgium.
E-mail address: aniko.lovik@kuleuven.be (A. Lovik).

http://dx.doi.org/10.1016/j.paid.2017.05.048
0191-8869/

time and space constraints. John and Srivastava (1999) developed the Big Five Inventory (BFI) to address these problems. Their questionnaire is available on the internet, making it useful for websurveys, and short, only consisting of 44 items, which can be rated in <15 min, while still reliable.

The analysis of questionnaires consisting of many items raises several methodological problems. To begin with, data reduction is needed. Questionnaires, even those composed of a dozen questions, ensue a complex structure of variables where covariances between responses have to be taken into account. Longer inventories are therefore often analysed using principal component or factor analysis which reduces the dimensionality by imposing a certain number of latent factors based on inter-item correlations. This is especially useful in personality research where the different personality traits can be reduced to personality factors. For example, personality traits expressed with adjectives such as “assertive”, “active”, “energetic”, “adventurous”, “outspoken” and “enthusiastic” can be collected in one term: Extraversion.

To be able to easily interpret these unobserved factors it is useful to assume that factors do not share common variance. In fact, the Big Five personality factors make this assumption: the five factors are independent, and this view is also implied by using an orthogonal (varimax) rotation which assumes uncorrelated factors. The varimax rotation, as its name suggests, maximises the sum of the variances of the squared (and normalised) factor loadings (Kaiser, 1959). Using this rotation when the factors have high correlations may result in a false interpretation of the results. A factor rotation not assuming uncorrelated factors, such as the direct oblimin rotation, may prove useful in such a situation.

Similarly, using ordinal, Likert-type data as continuous variables can affect the results of a factor analysis as calculations are based on the Pearson correlations between the variables. A simple alternative is to use polychoric correlations instead with the assumption that the variables are measured on an ordinal scale but the underlying parameter is continuous and normally distributed.

Another issue that may cause distortion in the results is related to the personal differences in filling-in questionnaires. Some participants tend to give extreme answers while others try to stay close to the centre of the scale and never use the two endpoints of a scale. At the same time some participants may prefer to give overwhelmingly positive or overwhelmingly negative responses. This tendency of agreement or disagreement with the items, independently of the item content has been known for some time in the literature (Jackson & Messick, 1958). While with a big sample size many of the individual differences are balanced out, some analyses are sensitive to these kinds of extreme responses and corrections are needed. Normalising the data by taking into account the individual differences in the centre and the variability of the answers may give a better understanding of the population. A dedicated normalisation, called the acquiescence index, has been presented by the inventors of the BFI in 2008 to account for this type of bias (Oliver, Naumann, & Soto, 2008) in this inventory. The method is also called ipsatisation.

Lastly, a general problem with survey data is the incompleteness of the collected data. The participants may forget or may not wish to fill in all answers resulting in a certain amount of missing data. An important aspect is the type of missingness, which can be divided into three categories according to Rubin (1976, 1987). Data is missing completely at random (MCAR) if the values that are missing are independent of both observed and unobserved outcomes. If the missingness can be fully accounted for by observed information, then the data are missing at random (MAR). Else, data are missing not at random (MNAR). While it is impossible to definitively determine for incomplete data to which category they belong, most analyses such as multiple imputation, assume that data are MAR or MCAR and result in biased estimates if the data is MNAR. To overcome this, sensitivity analysis is often argued for (Molenberghs & Kenward, 2007).

While some statistical methods can cope with a certain amount of (randomly) missing data, others cannot, which often results in the omission of observations from the analysis or even in the exclusion of an entire set of data. For example, because of one missing item we might need to exclude all answers of a participant from the analysis (this is called listwise deletion) even though >95% of the items are answered. Listwise deletion by definition reduces the power of the study. In case of the BFI an even bigger concern is the bias resulting from excluding participants with one or more missing items. Because we want to measure personality dimensions, participants who give themselves higher scores on the following items and can be somewhat careless (item 8), lazy (item 23), less organised (item 18) and/or easily distracted (item 43) may be more likely to miss a response which will be reflected on the same factor: Conscientiousness.

A possible solution is multiple imputation, a statistical method that allows the use of all observed data. A major difference between the analysis with listwise deletion and with multiple imputation can be an indicator of aforementioned bias.

The paper is organised as follows. The motivating dataset, the Divorce in Flanders study is first introduced. The psychometric properties and factor analysis of the original data using the standard varimax rotation is presented next, followed by results of the analysis with oblimin rotation and the factor analyses using polychoric correlations, the ipsatised data and the multiply imputed data, respectively. Finally, Discussion and concluding remarks are given.

2. Divorce in Flanders study

The dataset we use for analysis is a subsample from the ‘Divorce in Flanders’ (DiF) project, which contains a sample of marriages registered between 1971 and 2008 with oversampling of divorces (1/3 intact and 2/3 dissolved marriages at the sampling date) drawn from the Belgian National Register. Family members across three generations were surveyed during the original data collection, >10,000 people (Mortelmans et al., 2011). In this paper we use data from 4457 families, 7533 people in total (3362 mothers, 2920 fathers and 1251 children). We excluded new partners of the ex-spouses and parents of the selected sample but used the data from the new partners ($n = 1699$) for cross-validation. One of the main advantages of the data collection is the ability to assess, among others, the eventual patterns of matching personality traits between family members, predicting personality traits by studying the intergenerational transmission of personality, associating personality traits with fertility and personality traits with divorce.

As part of this study the personality of each participant was assessed with the validated Dutch version (Denissen, Geenen, van Aken, Gosling, & Potter, 2008) of the Big Five Inventory (BFI, John & Srivastava, 1999), a personality test which is a commonly used tool to assess personality measuring the five factors of personality (e.g. Goldberg, 1990; Widiger & Trull, 1997).

We considered individuals and not families as units, thereby neglecting the potential correlations between the responses of family members. Consequences of this decision will be taken up in the Discussion.

From each family one child (aged 10 or more) was selected randomly to participate but only children aged 14 or more were invited to fill in the BFI. If there was only one child above the age of 14 in the family, that child was automatically chosen. If the children in a specific family were all younger than 14 years at the time of the study they were not included in the study. However, the age of the children is of little concern as children give relatively stable responses to Big Five questionnaires from middle childhood (Asendorpf & van Aken, 2003). In a longitudinal study they found that all same-factor reliability measures at age 4–6 (teacher Q-scores) and at age 10 (parental Q-scores) were above 0.6. According to the Dutch BFI data, the factor structure does not change substantially with age (Denissen et al., 2008).