

Factorial structure of the indicators in Big Five Plus Inventory

Cristian Opariuc-Dan^{1,2,3}, Gabriela Nicuță³, and & Ticu Constantin³

¹ Ovidius University

Constanta

Romania

² Bucharest University

Bucharest

Romania

³ Al. I. Cuza University

Iassy

Romania

Author Note

The data in this paper comes form the Romanian version of BigFive Plus inventory(Constantin et al., 2019), being collected online through the PsihoProfile site

The authors thanks Prof. Dr. Constantin Ticu and the PsihoProfile team for making available over 14000 data used in this analysis.

The authors made the following contributions. Cristian Opariuc-Dan: Conceptualization, Writing - Original Draft Preparation, Writing - Data analysis; Gabriela Nicuță: Writing - Systematic review, Writing - Data analysis, Proofreading; Ticu Constantin: Data provider, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Cristian Opariuc-Dan, Ovidius University Campus, Aleea Universității, No. 2, Constanța, Romania. E-mail: copariuc@gmail.com

Abstract

TO BE DONE AT THE END

Keywords: BigFive, Confirmatory Factor Analysis, Reliability analysis

Word count: X

DRAFT

Factorial structure of the indicators in Big Five Plus Inventory

Methods

In this paper we aimed to conduct an exploratory study on the **analysis of how the theoretical Big Five model is supported by data**. The volume of data collected, even if it is very large and gives power to the study, does not constitute a representative sample because no sampling technique was used, but represents the results of tests performed with the Big Five Plus Personality Inventory (Constantin et al., 2019) during 2015-2020 on the population of Romania. The answers resulted either from an online administration using the PsihoProfile site, or from the classic administration, pencil-paper, followed by the introduction of the answers in the platform in order to automatically rate and generate assessment's protocols.

La nivel de bază, variabilele manifeste reprezintă cei 244 de itemi dihotomici ai Inventarului de Personalitate Big Five Plus, analiza intrând în domeniul analizelor factoriale confirmatorii folosindu-se indicatori categoriali, mai precis dihotomici, iar acest lucru implică o abordare specifică.

Modelul nu prevede existența unor indicatori încărcăți de mai mult de un singur factor latent, fiecare factor latent încărcând 8 indicatori, prin urmare și aceștia vor reprezenta variabile latente categoriale, deoarece o amplitudine de 8 puncte nu-i poate califica drept variabile continue.

Factorii latenți care încarcă cei 8 indicatori sunt la rândul lor încărcăți de una dintre cele 5 dimensiuni Big Five, rezultând un model de analiză factorială de ordin secund (Byrne, 2013) cu factori de ordin secund necorelați, aceștia fiind reprezentați după cum urmează:

- **Extraversion**, variabilă latentă exogenă care încarcă un număr de 6 variabile latente endogene: *Friendliness*, *Gregariousness*, *Assertiveness*, *Activity level*, *Excitement seeking* și *Cheerfulness*.

- **Agreeableness**, variabilă latentă exogenă care încarcă un număr de 6 variabile latente endogene: *Trust, Morality, Altruism, Cooperation, Modesty* și *Sympathy*.
- **Neuroticism**, variabilă latentă exogenă care încarcă un număr de 6 variabile latente endogene: *Anxiety, Anger, Depression, Timidity, Immoderation* și *Vulnerability*.
- **Conscientiousness**, variabilă latentă exogenă care încarcă un număr de 6 variabile latente endogene: *Self-efficacy, Orderliness, Dutifulness, Achievement-striving, Self-discipline* și *Cautiousness*.
- **Openness to experience**, variabilă latentă exogenă care încarcă un număr de 6 variabile latente endogene: *Imagination, Artistic interests, Emotionality, Adventurousness, Intellect* și *Liberalism*.

Participants

A total of 14706 protocols were collected (**n=14706**), administered to a number of 7907 males (53.80%) and 6799 females (6799%), most of them being *university graduates* (4014, representing 27.30%), followed by *high school graduated* (3388, representing 23%) and the persons with *master's degrees* (1764, representing 12%). The research group also contains graduates of *post-secondary schools* (1402, representing 9.50%) and graduates of *arts and crafts schools* (653, representing 4.40%), the other categories being much less represented.

Remarcăm, totuși, numărul mare al persoanelor care nu au precizat nivelul de studii (2813, reprezentând 19.10%), acestea urmând a fi eliminate din analizele ce presupun utilizarea acestei variabile.

Material

Instrumentul folosit este Inventarul de Personalitate Big Five Plus (Constantin et al., 2019), datele fiind colectate în perioada 2015-2020 (n=14706). Instrumentul, construit și validat

pe populația românească, permite evaluarea celor cinci meta-factori ai personalității (Ex-traversiune, Agreabilitate, Nevrozism, Conștiinciozitate și Deschidere), dar și a 30 de fațete subordonate acestora. Inventarul conține 240 de itemi dihotomici, câte 48 pentru evaluarea fiecăruia dintre cei cinci factori principali și câte opt pentru evaluarea celor 30 de fațete. Persoana este instruită să aleagă dintre două aserțiuni pe aceea care o descrie cel mai bine (e.g. *Atunci când sunt la o petrecere: (a) sunt în mijlocul acțiunii, înconjurat de ceilalți ; (b) prefer să stau de o parte și să observ.*). Cele două variante de răspuns reprezintă extreme ale dimensiunii psihologice vizate și se doresc a fi similare din perspectiva dezirabilității sociale. Timpul mediu de completare a chestionarului este de 35 de minute.

Procedure

Analiza va presupune studiul inițial al consistenței interne, pentru fiecare scală folosindu-se metoda α Cronbach (Cronbach, 1951) în vederea calculului coeficientului de consistență internă α (Guttman λ_3), dar și al coeficienților ierarhici de saturație ω (Zinbarg et al., 2007) deoarece deoarece instrumentul nu este unul unifactorial. În vederea analizei de structură internă și a adecvării la model unidimensional, vom testa, la nivelul fiecărei fațete, atât acest model, cât și două modele alternative, cu două și 3 sub-componente în vederea identificării unei structuri axiale (prezența unor scale cu sub-componente - lumpyness).

Vom testa apoi modelul în care factorul latent încarcă cei 8 indicatori, pentru fiecare dintre cele 6 fațete ale unei dimensiuni, și existența dimensiunii comune pentru toți cei 6 factori latenți.

În urma interpretării parametrilor și a diagnosticului modelelor acestea vor fi respecificate, identificându-se, dacă este cazul, modele explicative mai bune.

99 Data analysis

100 Toate analizele s-au realizat folosind limbajul R versiunea 4.1.2 și următoarele pachete: R
 101 (Version 4.1.2; R Core Team, 2021) and the R-packages *dplyr* (Version 1.0.7; Wickham et
 102 al., 2021), *epiDisplay* (Version 3.5.0.1; Chongsuvivatwong, 2018), *foreign* (Version 0.8.81; R
 103 Core Team, 2020), *lavaan* (Version 0.6.9; Rosseel, 2012), *MASS* (Version 7.3.54; Venables
 104 & Ripley, 2002a), *nnet* (Version 7.3.16; Venables & Ripley, 2002b), *nortest* (Version 1.0.4;
 105 Gross & Ligges, 2015), *papaja* (Version 0.1.0.9997; Aust & Barth, 2020), *psych* (Version
 106 2.1.9; Revelle, 2021), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *stargazer* (Version
 107 5.2.2; Hlavac, 2018), *survival* (Version 3.2.13; Terry M. Therneau & Patricia M. Grambsch,
 108 2000), and *tinylabels* (Version 0.2.1; Barth, 2021).

109 Results

110 **În prima fază** vom prezenta comparativ analizele autorilor privind consistența internă
 111 utilizându-se **metoda alpha Cronbach a fidelității tau-echivalente** (Cronbach, 1951),
 112 incluzând atât indicatorii globali, cât și cei defalcați în funcție de genul biologic și de modal-
 113 itatea de administrare. Analiza s-a realizat utilizându-se pachetul „**psych**” (Revelle, 2021),
 114 plecându-se de la condiția itemilor cumulativi, scorul total rezultând în urma adunării celor
 115 8 itemi (amplitudinea teoretică de 8 puncte) cu identificarea varianțelor negative ale itemilor
 116 și recodare automată și utilizând o reeșantionare neparametrică folosind 100 de eșantioane
 117 în vederea verificării stabilității parametrilor.

118 **În cea de-a doua fază** se vor calcula coeficienții ierarhici de saturație ω_h (Zinbarg et
 119 al., 2007), efectuându-se o analiză ierarhică a componentelor principale cu rotație oblică
 120 (*oblimin*) pe baza unei matrici de corelații tetrachorice cu inversarea automată a itemilor ce
 121 prezintă covarianțe negative cu factorul general, apoi aplicarea transformării Schmid Leiman
 122 asupra acesteia, testându-se astfel un număr de două modele în vederea identificării unei
 123 potențiale structuri „lumpyness”: **primul**, în care presupunem că itemii sunt încărcăți de

trei sub-componente ale factorului general, **al doilea**, în care vom pleca de la ipoteza în care itemii sunt încărcăți de două sub-componente ale factorului general, iar în acest caz vom trata trei situații: (1) încărcările factorului general vor fi redistribuite egal pe cele două subcomponente, (2) încărcările factorului general vor fi echivalate cu cele ale primei sub-componente și (3) încărcările factorului general vor fi echivalate cu ale celei de-a doua sub-componentă.

Confirmatory Factor Analysis

Analiza Inventarul BigFive Plus va fi realizată după un model clasic de analiză factorială confirmatorie cu factori de ordin secund necorelați, ce va fi analizat folosindu-se pachetul **lavaan** (Rosseel, 2012). Deoarece indicatorii sunt dihotomici, vom folosi *metoda estimării pragurilor categoriilor de răspuns* în care fiecare indicator devine încărcat în factor latent și determină un răspuns activ (1-Da) dacă depășește valoarea prag a sa, determinată de poziția itemului pe continuumul factorului latent. Drept estimator vom folosi unul din categoria metodelor bazate pe cele mai mici pătrate, mai precis „**WLSM**”, ce utilizează doar diagonala matricii ponderate „**W**” („**DWLS**”) și folosind, la testele statistici, mediile ajustate. În vederea testării potrivirii globale s-a folosit versiunea „Satorra Bentler” a testului χ^2 , ajustându-se mediile (**Skrondal-2005?**).

Extraversion

Varianța celor 48 de itemi dihotomici explicată de cei 6 factori latenți a căror varianță, mai apoi, va fi explicată de un factor de ordin doi, mai exact de **dimensiunea Extraversiune**, în urma analizei a generat o soluție stabilă după 29 de iterații, estimându-se un număr 61 parametri liberi, pentru o estimare validă fiind necesare minimum 610 observații, ideal 1220 observații, condiția volumului lotului de cercetare fiind îndeplinită.

Modelul global cu un singur factor latent de ordin doi nu este însă susținut în mod core-spunzător de date (*Robust $\chi^2_{(1115)}=55,751.85$, $p=0$*), testul de potrivire a modelului eșuând

și respingându-se ipoteza nulă H_0 : *Nu există nicio discrepanță statistic semnificativă între covarianțele stipulate la nivelul populației și covarianțele estimate de model.* Erorile de aproximare sunt însă acceptabile ($RMSEA=0.06$, $p=0$, $CI_{90\%}=0.06 - 0.06$), chiar dacă ipoteza nulă de potrivire H_0 : *Reziduurile standardizate dintre covarianțele rezultate din date și matricea ipotetică de covarianțe sunt nule* este respinsă, valoarea arătând eșecul testul de nepotrivire ($\epsilon > .10$) și reziduuri standardizate statistic semnificative între covarianțele rezultate din date și matricea ipotetică de covarianțe.

Indicele Tucker-Lewis de comparare cu modelul de bază ($Robust\ TLI=0.92$, $SRMR=0.08$) arată că modelul se îmbunătățește cu 92.07% în comparație cu modelul nul, în condițiile în care indicatorul standardizat al reziduurilor are o valoare ușor ridicată.

Din cauza unei probleme de identificare empirică și care a generat o matrice neinvertibilă, erorile standard ale estimării parametrilor nu au putut fi calculate, prin urmare nici testele statistice, așadar parametrii vor fi apreciați pe baza valorilor nestandardizate și standardizate.

Extraversiunea, ca factor latent de ordin doi, încarcă cel mai puternic factorul latent *Sociabilitate* ($B=0.78$, $\beta=0.95$, $R^2=0.90$) și cel mai slab factorul latent *Excitabilitate* ($B=0.49$, $\beta=0.55$, $R^2=0.30$), valorile varianței explicate și ale coeficientului de încărcare sugerând, de fapt, că *Excitabilitatea* nu reprezintă o fațetă a extraversiunii, cu atât mai mult cu cât cei 8 itemi sunt încărcăți foarte bine de aceasta.

În privința itemilor problematici, remarcăm itemul **I7** încărcat foarte slab de factorul **Asertivitate** ($B=0.25$, $\beta=0.19$, $R^2=0.04$), itemii **I28**, **I31** și **I37** încărcăți foarte slab de factorul **Activitate** (Itemul 28: $B=0.33$, $\beta=0.26$, $R^2=0.07$, Itemul 31: $B=0.46$, $\beta=0.36$, $R^2=0.13$ și Itemul 37: $B=0.48$, $\beta=0.37$, $R^2=0.14$), precum și itemul **I30** încărcat foarte slab de factorul **Veselie** ($B=0.49$, $\beta=0.40$, $R^2=0.16$)

Respecificând modelul prin eliminarea itemilor problematici și a factorului latent *Excitabilitate*, obținem o îmbunătățire a modelului în ce privește potrivirea globală ($\chi^2_{(584)}=30,643.11$,

p=0, RMSEA=0.06, p=0, CI_{90%}=0.06 - 0.06 față de $\chi^2_{(1115)}=55,751.85$, p=0, RMSEA=0.06, p=0, CI_{90%}=0.06 - 0.06), însă noul model tot nu este susținut în mod corespunzător de datele observate.

Singurul aspect remarcabil îl reprezintă creșterea capacității explicative față de modelul nul la 94.15% (TLI=0.94, SRMR=0.08) față de modelul inițial cu 92.07% (TLI=0.92, SRMR=0.08).

Discussion

References

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*.

<https://github.com/crsh/papaja>

Barth, M. (2021). *tinylabels: Lightweight variable labels*. [https://github.com/mariusbarth/](https://github.com/mariusbarth/tinylabels)

tinylabels

Byrne, B. M. (2013). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming, Second Edition* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203805534>

Chongsuvivatwong, V. (2018). *epiDisplay: Epidemiological data display package*.

<https://CRAN.R-project.org/package=epiDisplay>

Constantin, T., Gheorghiu, A., Căldare, L., Gervescu, A. E., Aiftincăi, A., Fodorea, A., Iliescu, M., Hojbotă, A. M., & Iordache, A. (2019). *Inventarul de personalitate Big Five Plus*. Psiho Profile.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>

Gross, J., & Ligges, U. (2015). *Nortest: Tests for normality*. <https://CRAN.R-project.org/package=nortest>

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. <https://CRAN.R-project.org/package=purrr>

Hlavac, M. (2018). *Stargazer: Well-formatted regression and summary statistics tables*. Central European Labour Studies Institute (CELSI). <https://CRAN.R-project.org/package=stargazer>

R Core Team. (2020). *Foreign: Read data stored by 'minitab', 's', 'SAS', 'SPSS', 'stata', 'sysstat', 'weka', 'dBase', ...* <https://CRAN.R-project.org/package=foreign>

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and person-*

ality research. Northwestern University. <https://CRAN.R-project.org/package=psych>

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://www.jstatsoft.org/v48/i02/>

Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending the Cox model*. Springer.

Venables, W. N., & Ripley, B. D. (2002a). *Modern applied statistics with s* (Fourth). Springer. <https://www.stats.ox.ac.uk/pub/MASS4/>

Venables, W. N., & Ripley, B. D. (2002b). *Modern applied statistics with s* (Fourth). Springer. <https://www.stats.ox.ac.uk/pub/MASS4/>

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr>

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating omega h for Structures Containing Two Group Factors: Perils and Prospects. *Applied Psychological Measurement*, 31(2), 135–157. <https://doi.org/10.1177/0146621606291558>

Table 1*Participants's gender*

	Frequency	Percent	Cum. percent
Male	7,907	53.800	53.800
Female	6,799	46.200	100
Total	14,706	100	100

Table 2*Participants's educational level*

	Frequency	%(NA+)	%(NA-)
Illiterate	11	0.100	0.100
Primary (4 years)	284	1.900	2.400
Gymnasium (8 years)	240	1.600	2
Arts and crafts school	653	4.400	5.500
Highschool (12 years)	3,388	23	28.500
Post graduated school	1,402	9.500	11.800
University (Bachelor level)	4,014	27.300	33.800
University (Master level)	1,764	12	14.800
Doctoral school	137	0.900	1.200
NA's	2,813	19.100	0
Total	14,706	100	100

Table 3*Loadings for Friendliness' latent trait*

Latent trait	Item	B	z	p	Beta
Friendliness	I1	0.85	NA	NA	0.67
Friendliness	I5	0.70	NA	NA	0.55
Friendliness	I10	1.06	NA	NA	0.83
Friendliness	I16	0.97	NA	NA	0.76
Friendliness	I18	0.80	NA	NA	0.63
Friendliness	I25	0.65	NA	NA	0.51
Friendliness	I34	0.65	NA	NA	0.51
Friendliness	I40	0.97	NA	NA	0.76

Table 4*Loadings for Gregariousness' latent trait*

Latent trait	Item	B	z	p	Beta
Gregariousness	I2	0.99	NA	NA	0.77
Gregariousness	I6	0.91	NA	NA	0.71
Gregariousness	I11	0.81	NA	NA	0.63
Gregariousness	I17	0.92	NA	NA	0.71
Gregariousness	I19	0.86	NA	NA	0.67
Gregariousness	I26	0.91	NA	NA	0.71
Gregariousness	I35	0.92	NA	NA	0.71
Gregariousness	I41	0.81	NA	NA	0.63

Table 5*Loadings for Assertiveness' latent trait*

Latent trait	Item	B	z	p	Beta
Assertiveness	I3	1.04	NA	NA	0.79
Assertiveness	I7	0.25	NA	NA	0.19
Assertiveness	I12	0.97	NA	NA	0.74
Assertiveness	I20	0.59	NA	NA	0.45
Assertiveness	I21	0.52	NA	NA	0.40
Assertiveness	I27	0.99	NA	NA	0.75
Assertiveness	I36	0.74	NA	NA	0.56
Assertiveness	I42	0.78	NA	NA	0.60

Table 6*Loadings for Activity's latent trait*

Latent trait	Item	B	z	p	Beta
Activity	I4	1.00	NA	NA	0.78
Activity	I13	0.56	NA	NA	0.43
Activity	I22	0.98	NA	NA	0.76
Activity	I28	0.33	NA	NA	0.26
Activity	I31	0.46	NA	NA	0.36
Activity	I37	0.48	NA	NA	0.37
Activity	I43	0.99	NA	NA	0.77
Activity	I44	0.84	NA	NA	0.65

Table 7*Loadings for Excitement's latent trait*

Latent trait	Item	B	z	p	Beta
Excitement	I8	0.62	NA	NA	0.52
Excitement	I14	0.95	NA	NA	0.81
Excitement	I23	0.92	NA	NA	0.78
Excitement	I29	0.80	NA	NA	0.68
Excitement	I32	0.83	NA	NA	0.71
Excitement	I38	0.81	NA	NA	0.69
Excitement	I45	0.69	NA	NA	0.58
Excitement	I47	0.87	NA	NA	0.74

Table 8*Loadings for Cheerfulness' latent trait*

Latent trait	Item	B	z	p	Beta
Cheerfulness	I9	0.90	NA	NA	0.73
Cheerfulness	I15	0.66	NA	NA	0.54
Cheerfulness	I24	1.09	NA	NA	0.89
Cheerfulness	I30	0.49	NA	NA	0.40
Cheerfulness	I33	0.75	NA	NA	0.61
Cheerfulness	I39	0.70	NA	NA	0.57
Cheerfulness	I46	0.56	NA	NA	0.46
Cheerfulness	I48	0.69	NA	NA	0.56