

# Assignment 2

SEM 1: Confirmatory Factor Analysis

Please hand in a .pdf file containing your report and the .csv file that has been generated. The deadline of this assignment is Thursday April 19 15:00.

## Assignment

In R, run the following code in R (replace . . . with your student number, or your birthday (YYYYMMDD) if you do not have a student number):

```
library("install.packages")
library("psych")
studentNumber <- ...
set.seed(studentNumber)
data(bfi)
noNA <- na.omit(bfi)
Data <- noNA[sample(1:nrow(noNA), 500), 1:25]
write.csv(Data,paste0("bfi_",studentNumber,".csv"), row.names = FALSE)
```

This code created an object `Data`, which is also stored as a csv file in your working directory. If you do not use R, let someone run this code for you and send you the csv file (do use your own student number). This dataset contains your personal  $n = 500$  subset of the `bfi` (Big Five Inventory) data contained in the `psych` package. More information on this dataset can be obtained by running:

```
?bfi
```

The questions are designed to measure five central personality traits: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness. The following table gives the item descriptions:

Item label	Item description	Trait
A1	Am indifferent to the feelings of others	Agreeableness
A2	Inquire about others' well-being	Agreeableness
A3	Know how to comfort others	Agreeableness
A4	Love children	Agreeableness
A5	Make people feel at ease	Agreeableness
C1	Am exacting in my work	Conscientiousness
C2	Continue until everything is perfect	Conscientiousness
C3	Do things according to a plan	Conscientiousness
C4	Do things in a half-way manner	Conscientiousness
C5	Waste my time	Conscientiousness
E1	Don't talk a lot	Extraversion
E2	Find it difficult to approach others	Extraversion
E3	Know how to captivate people	Extraversion
E4	Make friends easily	Extraversion
E5	Take charge	Extraversion
N1	Get angry easily	Neuroticism
N2	Get irritated easily	Neuroticism
N3	Have frequent mood swings	Neuroticism
N4	Often feel blue	Neuroticism
N5	Panic easily	Neuroticism
O1	Am full of ideas	Openness
O2	Avoid difficult reading material	Openness
O3	Carry the conversation to a higher level	Openness
O4	Spend time reflecting on things	Openness
O5	Will not probe deeply into a subject	Openness

Your assignment is to investigate the Big 5 personality structure, as if you collected your subset of the data yourself, and write a short scientific report (no more than 1000 words) on your findings. Include at least the following:

- Fit the theoretical 5-factor model (each item only loading on the intended traits, no residual correlations and all factors correlated). Report the Lavaan model used (may be generated using Onyx) and include a clear path diagram. Judge the fit of the CFA model.
- Perform exploratory model modifications until your model fits *acceptably* well. You may decide for yourself what “acceptably well” means (define this in your methods section). Report the Lavaan model of your final model and a clear path diagram.
  - You may employ any strategy you want. For instance, you can use modification indices to stepwise add parameters (do mention this), or you may think about how to extend the model before fitting by investigating the items (e.g., you might expect a cross-loading before analyzing)
  - Include a table of all consecutive models you fit, including AIC, BIC and  $\chi^2$  difference tests with previous model (if nested)
- Discuss the plausibility of the added model modifications (e.g., did you add cross-loadings that did not make sense from a theoretical perspective?).

Your report should read as a paper (introduction, methods, results, discussion), and will be graded as a writing assignment in addition to how well the analyses are performed and described. You may use semPlot or Onyx to make your path diagrams, but make sure the path diagrams are clear and interpretable. In semPlot, at least use:

```
semPaths(..., style = "lisrel", reorder = FALSE)
```

Add any graphical arguments explained in `?semPaths` you want. In addition to your report, also send the data you used. Note, every student will have their own dataset to work on, and as a result every student can get different results! I will show the path diagrams you generated in the final lecture to investigate how much the results differ.

## Challenge Question

If you want an extra challenge (bonus credit), you may try state-of-the-art exploratory model modification options. For example, you can try the *regsem* package (Jacobucci, Grimm, & McArdle, 2016; Jacobucci, 2017) for regularized CFA models (using LASSO regularization to add cross-loadings) or the *lvnet* package (Epskamp, Rhemtulla, & Borsboom, 2017) for residual network modeling (using LASSO regularization to add a residual network structure, will be explained in the last lecture of SEM2).

## References

- Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 1–24. Retrieved from <http://dx.doi.org/10.1007/s11336-017-9557-x> doi: 10.1007/s11336-017-9557-x
- Jacobucci, R. (2017). regsem: Regularized structural equation modeling. *arXiv preprint arXiv:1703.08489*.
- Jacobucci, R., Grimm, K. J., & McArdle, J. J. (2016). Regularized structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, 23(4), 555–566.