**Lab Day 2a: EFA and CFA**

Today’s lab meeting consists of three exercises:

In exercise 1, you replicate what was done in the lecture with respect to the basics of factor analysis. Exercises 2 and 3 are additional exercises on EFA and CFA, respectively.

Practical information:

* All the data and other files for these exercises can be found at the LLL platform.

Make sure to unzip the files. The folder containing these files will be your working directory.

* Solutions to the exercises can be found in the Solutions folder.

We provided R scripts for doing each of the exercises with lavaan.

***Exercise 1: Replicate Sapi results from the lecture***

The file **Sapi.txt** contains the data that was used to conduct the EFA and CFA as presented in the lecture. Try to replicate the following input files and/or results of today’s presentation and see what happens to the results.

Note: To prevent issues, make sure you specify -999 as missing data.

a) ***Reflective (confirmatory) Factor Model***

* Inspect the correlations for the items of interest: Q77 Q84 Q170 Q196.

*Why would you inspect these before doing a CFA?*

* Replicate the results for the extraversion model presented in the lecture (slides).

One can use the function lavaan() but also the function cfa() for this, the solutions files use the latter one.

b) ***Scaling your latent variable***

* Is the latent variable scaled via the variance of the factor (reference group scaling), or via the factor loadings (marker variable scaling)?

Use the function lavInspect() to obtain insight.

* Adjust the model so that you scale in the other way.

That is, if part a) was scaled by fixing a factor loading to 1, then free this factor loading and fix the factor variance to 1; If part a was scaled by fixing the factor variance to 1, free the factor variance, and fix one of the factor loadings to 1.

* Bonus: Scale via the factor loadings again, but now fix the factor loading of Q196 to 1, and free the factor loading that was previously fixed to 1.

Based on your results, fill out the table below.

*What do you notice?*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Scaling Method** | Chi-square value, df, p-value | Factor loading Q77 | Factor loading  Q84 | Factor loading  Q170 | Factor loading  Q196 | Factor variance |
| Reference Group |  |  |  |  |  |  |
| Marker Variable |  |  |  |  |  |  |
| Marker Variable Bonus |  |  |  |  |  |  |

c) ***Model fit***

Inspect the model fit. Focus on the CFI, TLI, RMSEA, and SRMR

One can use summary(<fit >, fit.measures=TRUE) and/or fitMeasures() and/or fitMeasures(<fit >, c("cfi", "tli", "rmsea","srmr")).

d) ***Test equality of factor loadings***

Based on the factor loadings, items do not seem to be equally reflective of Extraversion, except for Q84 and Q196. To improve the fit of the model, one can test whether these factor loadings are equal (resulting in estimating less parameters). One can use a Wald test to test whether two factor loadings can be considered equal. Test whether the factor loadings of Q84 and Q196 are the same.

For a lavaan object (with label names!), one can do this with lavTestWald(fit, constraints = ‘<insert equality restriction>’).

*What does the test imply?*

e) ***EFA versus CFA***

Specify and run the two-factor exploratory factor model (EFA) and two-factor CFA, as was shown in the lecture (slides).

***Exercise 2: EFA***

Use the data file **popular\_factor.txt**. Make sure to denote that 99 represents the missing values. Run an EFA that provides you with a 1-factor and 2-factor solution.

Interpret the results for the 2-factor solution.

Which factor model do you prefer? Note: you can use for example the anova() function for this.

***Exercise 3: CFA***

Use again **popular\_factor.txt**.Now, run a CFA with 1 factor (‘anti’) and another model with 2 factors (‘covert’ and ‘overt’). In both models, scale by fixing the factor variance, rather than by fixing a loading.

*Which model is to be preferred? What can you say about the model fit based on the AIC and BIC?*

*What is the correlation between the two factors?*