Confirmatory Factor Analysis

Statsomat.com

23 December 2020

Warning: The automatic computation and interpretation delivered by the Statsomat should not completely replace the classical, made by humans graphical exploratory data analysis and statistical analysis. There may be data cases for which the Statsomat does not deliver the most optimal solution or output interpretation.

Basic Information

Automatic statistics for the file:

File
neuroticism.csv

Your selection for the encoding: Auto
Your selection for the decimal character: Auto
Observations (rows with at least one non-missing value): 250
Variables (columns with at least one non-missing value): 8
Variables considered continuous: 8

Variables considered continuous
N1
N2
N3
N4
E1
E2
E3
E4

Numerical variables considered binary or ordinal: 0
Character variables considered binary: 0
Character variables considered nominal and transformed to binary: 0

Model Syntax

The following table describes the applied model equations in lavaan model syntax, either as entered by you in the text area (denoted by User=1) or established internally (User=0). The last column numbers the free parameters which are estimated.

Left hand side	Operator	Right hand side	User	Free parameter
NEUROT	=~	N1	1	0
NEUROT	=~	N2	1	1
NEUROT	=~	N3	1	2
NEUROT	=~	N4	1	3
EXTRAV	=~	E1	1	0
EXTRAV	=~	E2	1	4
EXTRAV	=~	E3	1	5
EXTRAV	=~	E4	1	6
N1	~~	N1	0	7
N2	~~	N2	0	8
N3	~~	N3	0	9
N4	~~	N4	0	10
E1	~~	E1	0	11
E2	~~	E2	0	12
E3	~~	E3	0	13
E4	~~	E4	0	14
NEUROT	~~	NEUROT	0	15
EXTRAV	~~	EXTRAV	0	16
NEUROT	~~	EXTRAV	0	17
N1	~1		0	18
N2	~1		0	19
N3	~1		0	20
N4	~1		0	21
E1	~1		0	22
E2	~1		0	23
E3	~1		0	24
E4	~1		0	25
NEUROT	~1		0	0
EXTRAV	~1		0	0

Assumptions

Open issue

Model Settings

Outputs

Model Fit Summary

```
lavaan 0.6-7 ended normally after 59 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of free parameters                         25

  Number of observations                           250
  Number of missing patterns                         1

Model Test User Model:

  Test statistic                                13.285
  Degrees of freedom                                19
  P-value (Chi-square)                           0.824

Model Test Baseline Model:

  Test statistic                              1253.791
  Degrees of freedom                                28
  P-value                                        0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000
  Tucker-Lewis Index (TLI)                       1.007

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -5748.501
  Loglikelihood unrestricted model (H1)      -5741.858

  Akaike (AIC)                               11547.002
  Bayesian (BIC)                             11635.038
  Sample-size adjusted Bayesian (BIC)        11555.786

Root Mean Square Error of Approximation:

  RMSEA                                          0.000
  90 Percent confidence interval - lower         0.000
  90 Percent confidence interval - upper         0.034
  P-value RMSEA <= 0.05                          0.990
```

```
Standardized Root Mean Square Residual:

  SRMR                                           0.018

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Observed
  Observed information based on                Hessian

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  NEUROT =~
    N1                1.000
    N2                0.942    0.052   18.146    0.000
    N3                1.071    0.061   17.592    0.000
    N4                0.997    0.052   19.217    0.000
  EXTRAV =~
    E1                1.000
    E2                1.074    0.078   13.713    0.000
    E3                0.935    0.072   12.921    0.000
    E4                0.814    0.073   11.111    0.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  NEUROT ~~
    EXTRAV          -10.512    1.932   -5.442    0.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .N1                0.000    0.360    0.000    1.000
   .N2                0.000    0.353    0.000    1.000
   .N3               -0.000    0.404   -0.000    1.000
   .N4               -0.000    0.360   -0.000    1.000
   .E1               -0.000    0.379   -0.000    1.000
   .E2               -0.000    0.391   -0.000    1.000
   .E3               -0.000    0.360   -0.000    1.000
   .E4                0.000    0.353    0.000    1.000
    NEUROT            0.000
    EXTRAV            0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .N1                7.025    0.914    7.688    0.000
   .N2                8.746    1.002    8.731    0.000
   .N3               11.760    1.336    8.799    0.000
   .N4                7.188    0.921    7.802    0.000
```

```
   .E1               12.802    1.580    8.104    0.000
   .E2               11.671    1.603    7.283    0.000
   .E3               12.192    1.464    8.327    0.000
   .E4               15.972    1.668    9.574    0.000
    NEUROT           25.335    2.902    8.730    0.000
    EXTRAV           23.054    3.187    7.233    0.000
```

Completely Standardized Parameter Estimates

```
Latent Variables:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  NEUROT =~
    N1               0.885    0.018   49.253    0.000    0.850    0.920
    N2               0.849    0.021   39.868    0.000    0.807    0.890
    N3               0.844    0.022   38.674    0.000    0.801    0.886
    N4               0.882    0.018   48.478    0.000    0.846    0.918
  EXTRAV =~
    E1               0.802    0.029   27.214    0.000    0.744    0.860
    E2               0.834    0.027   30.766    0.000    0.781    0.887
    E3               0.789    0.031   25.839    0.000    0.730    0.849
    E4               0.699    0.038   18.451    0.000    0.625    0.773

Covariances:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  NEUROT ~~
    EXTRAV          -0.435    0.059   -7.397    0.000   -0.550   -0.320

Intercepts:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .N1               0.000    0.063    0.000    1.000   -0.124    0.124
   .N2               0.000    0.063    0.000    1.000   -0.124    0.124
   .N3              -0.000    0.063   -0.000    1.000   -0.124    0.124
   .N4              -0.000    0.063   -0.000    1.000   -0.124    0.124
   .E1              -0.000    0.063   -0.000    1.000   -0.124    0.124
   .E2               0.000    0.063    0.000    1.000   -0.124    0.124
   .E3              -0.000    0.063   -0.000    1.000   -0.124    0.124
   .E4               0.000    0.063    0.000    1.000   -0.124    0.124
    NEUROT           0.000                               0.000    0.000
    EXTRAV           0.000                               0.000    0.000

Variances:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .N1               0.217    0.032    6.829    0.000    0.155    0.279
   .N2               0.280    0.036    7.753    0.000    0.209    0.351
   .N3               0.288    0.037    7.832    0.000    0.216    0.360
   .N4               0.222    0.032    6.921    0.000    0.159    0.285
   .E1               0.357    0.047    7.556    0.000    0.264    0.450
```

```
 .E2                0.305    0.045    6.746    0.000    0.216    0.393
 .E3                0.377    0.048    7.810    0.000    0.282    0.471
 .E4                0.511    0.053    9.654    0.000    0.408    0.615
  NEUROT            1.000                               1.000    1.000
  EXTRAV            1.000                               1.000    1.000
```

Communality

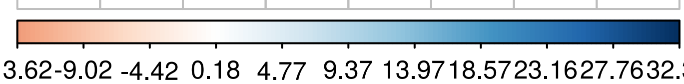
Table 1: Communality

in which order is it?

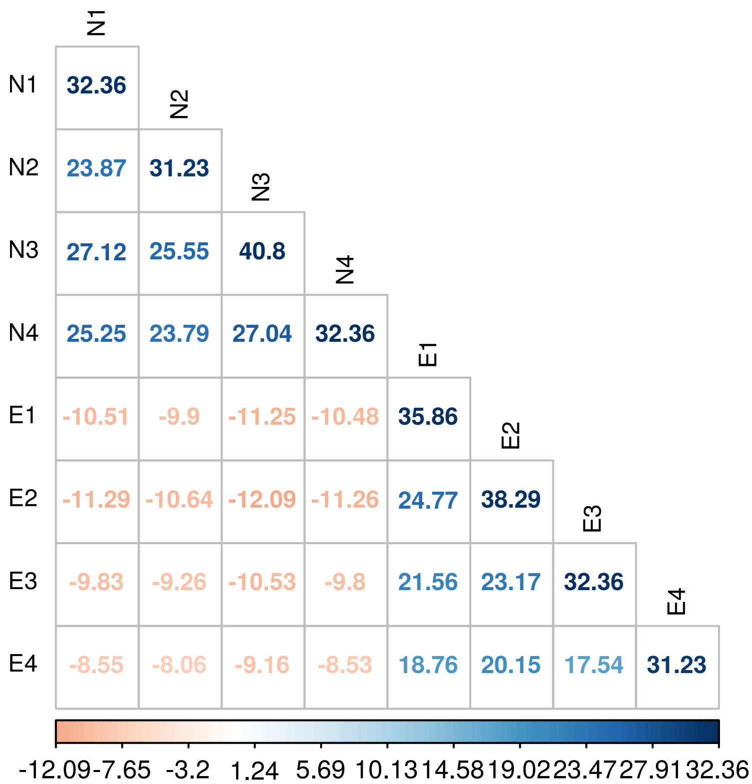
Variable	Communality
N1	0.78
N2	0.72
N3	0.71
N4	0.78
E1	0.64
E2	0.70
E3	0.62
E4	0.49

Observed Covariance Matrix

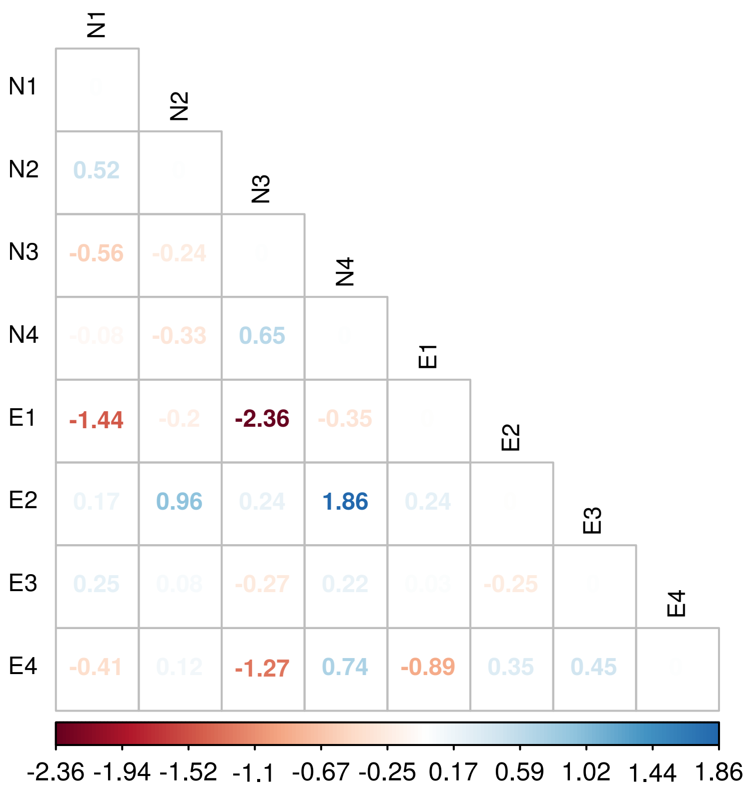
	N1	N2	N3	N4	E1	E2	E3	E4
N1	32.36							
N2	24.38	31.23						
N3	26.56	25.31	40.8					
N4	25.18	23.46	27.69	32.36				
E1	-11.96	-10.11	-13.62	-10.83	35.86			
E2	-11.12	-9.68	-11.86	-9.4	25.01	38.29		
E3	-9.58	-9.19	-10.79	-9.58	21.6	22.91	32.36	
E4	-8.97	-7.93	-10.42	-7.79	17.87	20.51	17.99	31.23



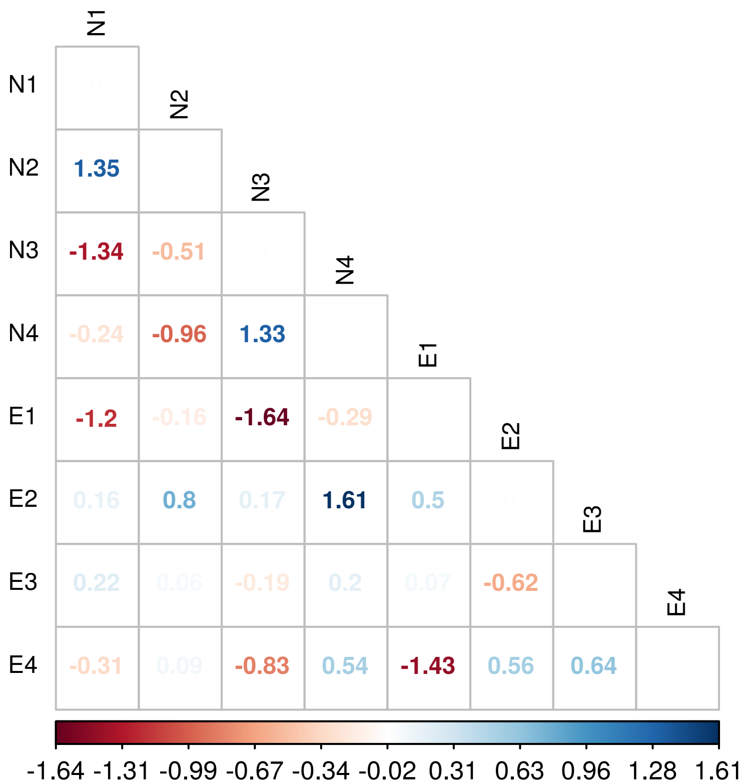
Fitted Covariance Matrix



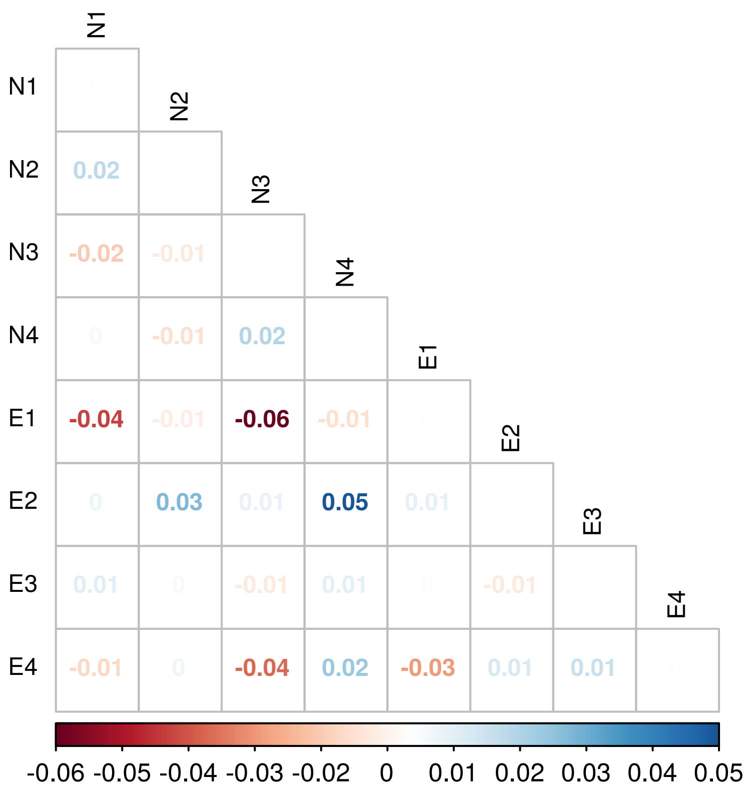
Residual Covariance Matrix



Standardized Residual Matrix



Residual Correlation Matrix



Modification Indices

define possible operators
in the statsomat interface

Table 2: Modification Indices With Respect To (Residual) Correlation

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
N3	~~	N4	3.507	1.820	0.1	0.051	(i)
N1	~~	N2	3.227	1.538	0.1	0.052	(i)
N1	~~	N3	2.747	-1.615	0.1	0.051	(i)
E1	~~	E4	1.983	-1.827	0.1	0.051	(i)
N4	~~	E2	1.276	0.906	0.1	0.052	(i)
N2	~~	N4	1.253	-0.956	0.1	0.052	(i)
N3	~~	E1	1.101	-1.028	0.1	0.051	(i)
N1	~~	E3	0.919	0.740	0.1	0.052	(i)
N4	~~	E3	0.838	-0.711	0.1	0.052	(i)
N2	~~	E1	0.658	0.689	0.1	0.052	(i)
N3	~~	E4	0.587	-0.785	0.1	0.051	(i)
E3	~~	E4	0.503	0.872	0.1	0.051	(i)
N1	~~	E2	0.500	-0.563	0.1	0.052	(i)
N4	~~	E4	0.480	0.584	0.1	0.052	(i)
E2	~~	E3	0.463	-1.037	0.1	0.050	(i)
E1	~~	E2	0.428	1.071	0.1	0.050	(i)
E2	~~	E4	0.412	0.873	0.1	0.051	(i)
N2	~~	N3	0.311	-0.525	0.1	0.051	(i)
N1	~~	E1	0.255	-0.405	0.1	0.052	(i)
N3	~~	E3	0.254	0.476	0.1	0.051	(i)
N2	~~	E3	0.222	-0.385	0.1	0.052	(i)
N1	~~	N4	0.121	-0.317	0.1	0.051	(i)
N3	~~	E2	0.107	0.319	0.1	0.051	(i)
N1	~~	E4	0.016	-0.107	0.1	0.052	(i)
E1	~~	E3	0.006	0.107	0.1	0.051	(i)
N2	~~	E2	0.004	-0.056	0.1	0.052	(i)
N2	~~	E4	0.002	0.041	0.1	0.051	(i)
N4	~~	E1	0.000	-0.005	0.1	0.052	(i)

Table 3: Modification Indices With Respect To Factor Loadings

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
NEUROT	=~	E1	1.328	-0.073	0.4	1	(nm)
NEUROT	=~	E2	1.121	0.068	0.4	1	(nm)
EXTRAV	=~	N4	1.067	0.053	0.4	1	(nm)
EXTRAV	=~	N3	0.904	-0.059	0.4	1	(nm)
EXTRAV	=~	N1	0.273	-0.027	0.4	1	(nm)
EXTRAV	=~	N2	0.143	0.020	0.4	1	(nm)
NEUROT	=~	E3	0.014	0.007	0.4	1	(nm)
NEUROT	=~	E4	0.013	-0.007	0.4	1	(nm)

Interpretation

Goodness of Fit Indices

We consider some of the model fit indices from the Model Fit Summary section to check the goodness-of-fit of the model. To decide for an acceptable or non-acceptable model, we apply thresholds considered in the References: [@brown], [@kline].

Model Test User Model
The degrees of freedom are calculated as the number of known parameters minus the number of free parameters: 44 - 25 = 19 . The 19 degrees of freedom indicate an over-identified model, fact which basically enables further analysis and interpretation.

The test statistic with the value 13.285 is called the Chi-square model fit index and represents the difference between summaries of the model-implied covariance matrix and the observed covariance matrix which is hypothesized and desirable to be zero. In general, if the p-value is larger than 0.05 then the test is not statistically significant at 5 % error, the hypothesis cannot be rejected, which would be in favour of the model.
In our case, the p-value is 0.824 suggesting an acceptable model fit.

Model Test Baseline Model
The test statistic with the value 1253.791 represents the difference between summaries of the baseline model (an alternative model-implied covariance matrix having zero covariances, i.e. a worst fitting model assuming independent variables) and the observed covariance matrix. The p-value of the test of a zero difference is <0.001 suggesting that the baseline model does not fit good to the data. This result is used indirectly in the construction of other model fit indices.

Root Mean Square Error of Approximation:
The Root Mean Square Error of Approximation (RMSEA) is a fit index based on the chi-square test statistic, which corrects for parsimony, i.e. overly complex models are penalized. RMSEA can be greater or equal than zero, with values close to zero suggesting an acceptable model fit.
In our case, the RMSEA is 0. The upper bound of the 90% confidence interval of the RMSEA is 0.034 and smaller than the threshold value 0.05, suggesting an excellent model fit.

Standardized Root Mean Square Residual:
The Standardized Root Mean Square Residual (SRMR) is a fit index derived from the residual correlation matrix with a range between zero and one with values close to zero suggesting an acceptable model fit.
In our case, the SRMR is 0.02 which is smaller than the threshold value 0.1 suggesting an acceptable model fit.

User Model versus Baseline Model
The Comparative Fit Index (CFI), evaluates the fit of the the model in relation to the worst-fitting baseline model described above. It ranges between zero and one, with values close to one suggesting good models (in the sense of departure from the baseline model).
In our case, the CFI is 1 which is greater or equal than the threshold value 0.95, suggesting a good model fit.

Similarly to the CFI, the Tucker-Lewis Index (TLI) evaluates the fit of the model in relation to the worst-fitting baseline model described above. Moreover, overly complex models are penalized. Values can range outside zero and one but the index is interpreted similarly to the CFI.
In our case, the TLI is 1.007 which is greater or equal than the threshold value 0.95, suggesting a good model fit.

Summary of the Goodness of Fit Indices

The Chi-square model fit index, the RMSEA and the TLI suggest an acceptable model fit. We tentatively assume an acceptable model fit and verify this assertion by considering further metrics.

Residuals

We analyze the residual matrices from the Outputs chapter. The residual covariance matrix represents the difference between the observed covariance matrix and the fitted model-implied covariance matrix. Large absolute values indicate local areas of misfit. However, the residuals are affected by the raw metric and are difficult to interpret more precisely.
A better interpretation allows the standardized residual matrix (residuals divided by their estimated asymptotic standard error) and the residual correlation matrix.
There are no variable pairs with standardized residuals which are larger or equal than the considered threshold 2.58 [@brown] or correlation residuals which are larger or equal than the considered threshold 0.1 [@kline]. Therefore, no relationships among the variables are substantially underestimated by the model. There are no variable pairs with standardized residuals which are smaller or equal than the considered threshold -2.58 [@brown] or correlation residuals which are smaller or equal than the considered threshold -0.1 [@kline]. Therefore, no relationships among the variables are substantially overestimated by the model.

Modification Indices

In the interpretation of the modification indices we rely mostly on [@brown] and [@mi]. We cite from [@brown]: “The modification index reflects an approximation of how much the overall model Chi² will decrease if the fixed or constrained parameter is freely estimated.” In other words, if adding a line with a high modification index to the model, i.e. if adding a parameter, the overall goodness-of-fit may be improved. Nevertheless, this should be done only under certain conditions, described in the sequel.
We consider only modification indices greater or equal than 3.84 (which are statistically significant at 5% type I error). Next, we search only for modification indices which achieve a power of minimum 75% in detecting a (relevant) misspecification of at least 0.1 for (residual) correlations, respectively 0.4 for factor loadings. These are characterized in the decision column by the label “epc:m”.
We remark that these conditions are not fulfilled for modification indices with respect to (residual) correlations. Therefore, there exist no significant and relevant modification indices with respect to (residual) correlations.
We remark that these conditions are not fulfilled for modification indices with respect to factor loadings. Therefore, there exist no significant and relevant modification indices with respect to factor loadings.

Parameter Estimates

Factor Loadings

We remark that the completely standardized factor loadings (section “Completely Standardized Parameter Estimates”) are all statistically significant at 5% type I error. Moreover, in absolute value they are all greater than 0.4. This cutoff-value is considered in some CFA research areas a magnitude that is substantively meaningful [@brown]. Please consider also cutoff-values from your particular research area when interpreting the factor loadings. We summarize the interpretation of the completely standardized factor loadings in the next table:

Table 4: Check Completely Standardized Factor Loadings

Latent Variable	Observed Variable	Factor Loading[1]	P-Value	Significant?[2]	Relevance[3]	Direction	Check
NEUROT	N1	0.88	<0.001	Yes	***	Ok	Ok
NEUROT	N2	0.85	<0.001	Yes	***	Ok	Ok
NEUROT	N3	0.84	<0.001	Yes	***	Ok	Ok
NEUROT	N4	0.88	<0.001	Yes	***	Ok	Ok
EXTRAV	E1	0.80	<0.001	Yes	***	Ok	Ok
EXTRAV	E2	0.83	<0.001	Yes	***	Ok	Ok
EXTRAV	E3	0.79	<0.001	Yes	**	Ok	Ok
EXTRAV	E4	0.70	<0.001	Yes	**	Ok	Ok

[1] The completely standardized factor loading can be interpreted as the correlation with the factor.
[2] 5% type I error.
[3] Stars correspond to factor loadings cutoff-values: 0.4, 0.6, 0.8.
[4] The observed variable is probably not related to latent factor.
[5] Uncertain. The evidence is insufficient or the model is misspecified.
[6] Uncertain. Significant but small(er) effect size.

Next, we inspect the unstandardized factor loadings (section "Model Fit Summary"). We remark that the unstandardized factor loadings are all statistically significant at 5% type I error. Therefore, the significance test results for standardized and unstandardized factor loadings coincide (for non-marker variables). We summarize the interpretation of the unstandardized factor loadings in the next table(s):

Table 5: Check Unstandardized Factor Loadings

Latent Variable	Observed Variable	Factor Loading	P-Value	Significant?[1]	Direction
NEUROT	N1	1.00			Ok
NEUROT	N2	0.94	<0.001	Yes	Ok
NEUROT	N3	1.07	<0.001	Yes	Ok
NEUROT	N4	1.00	<0.001	Yes	Ok
EXTRAV	E1	1.00			Ok
EXTRAV	E2	1.07	<0.001	Yes	Ok
EXTRAV	E3	0.94	<0.001	Yes	Ok
EXTRAV	E4	0.81	<0.001	Yes	Ok

[1] 5% type I error.

Table 6: Interpretation of Unstandardized Factor Loadings

Interpretation of Unstandardized Factor Loadings
N1 is marker variable for NEUROT
A 1-unit increase in NEUROT leads to a 0.94 -unit increase in the N2
A 1-unit increase in NEUROT leads to a 1.07 -unit increase in the N3
A 1-unit increase in NEUROT leads to a 1.00 -unit increase in the N4
E1 is marker variable for EXTRAV
A 1-unit increase in EXTRAV leads to a 1.07 -unit increase in the E2
A 1-unit increase in EXTRAV leads to a 0.94 -unit increase in the E3
A 1-unit increase in EXTRAV leads to a 0.81 -unit increase in the E4

Factor Correlations

As noted by [@brown], “the interpretability of the size and statistical significance of factor intercorrelations depends on the specific research context.”Though, the largest estimated factor intercorrelation within the section “Completely Standardized Parameter Estimates” is -0.43 which we regard as a proof of a reasonable discriminant validity.

Factor Reliability

Error Variances

Table 7: Completely Standardized Error Variances and Communality

Observed Variable	Error Variance[1]	Communality[23]	P-Value	Significant Error Variance?[4]
E1	0.36	0.64	<0.001	Yes
E2	0.30	0.70	<0.001	Yes
E3	0.38	0.62	<0.001	Yes
E4	0.51	0.49	<0.001	Yes
N1	0.22	0.78	<0.001	Yes
N2	0.28	0.72	<0.001	Yes
N3	0.29	0.71	<0.001	Yes
N4	0.22	0.78	<0.001	Yes

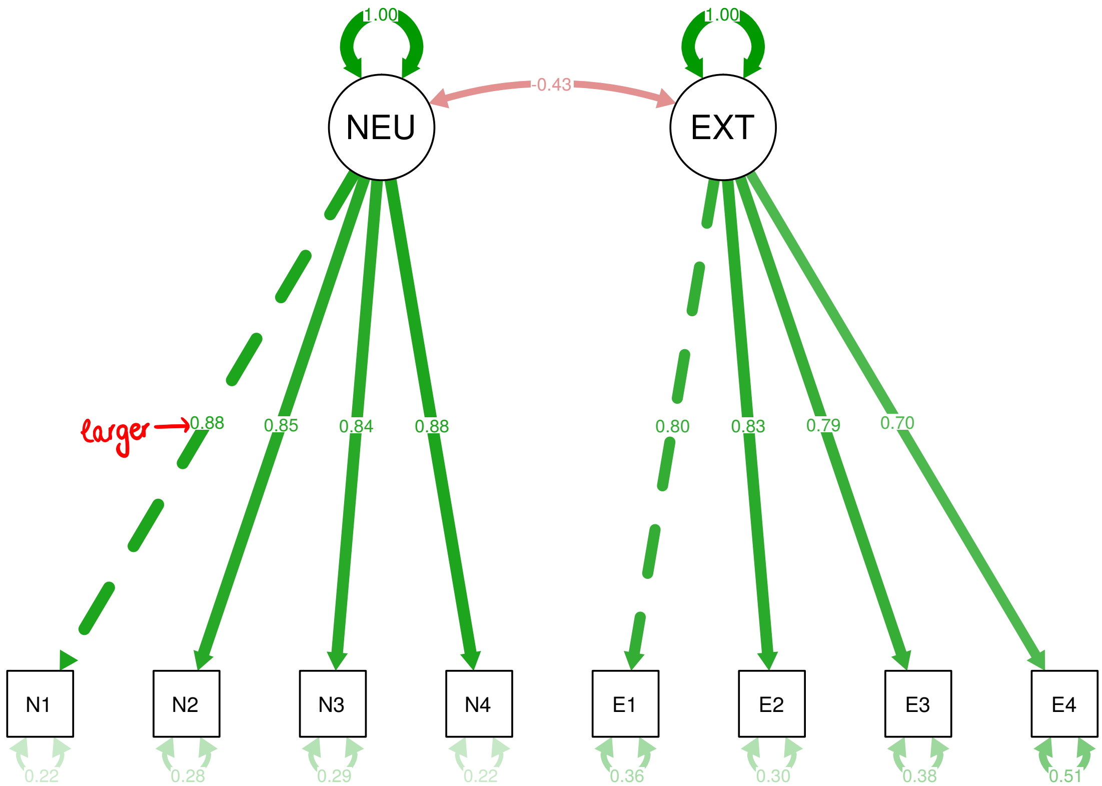
[1] Can be interpreted as proportion of unexplained variance by the latent factor(s) (%).
[2] Corresponds to the squared factor loading.
[3] Can be interpreted as proportion of explained variance by the latent factor(s) (%).
[4] 5% type I error. Typically significant since a large portion of variance is not explained by the latent variable.

Intercepts

In case of missing values and estimation via FIML, a meanstructure i.e. the intercepts of the observed variables are added to the model. The means of the latent factors are fixed to zero. Therefore, the estimated intercepts within the section “Model Fit Summary” are just the means of the observed variables.

+ add: Is there a better model? → which?

Path Diagram



References