Confirmatory Factor Analysis

Statsomat.com

07 January 2021

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
case18.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): 9
Variables considered continuous: 9

Variables considered continuous
PARI
PARC
PARO
SZTI
SZTC
SZTO
SZDI
SZDC
SZDO

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
PARANOID	=~	PARI	1	0
PARANOID	=~	PARC	1	1
PARANOID	=~	PARO	1	2
SCHIZOTYPAL	=~	SZTI	1	0
SCHIZOTYPAL	=~	SZTC	1	3
SCHIZOTYPAL	=~	SZTO	1	4
SCHIZOID	=~	SZDI	1	0
SCHIZOID	=~	SZDC	1	5
SCHIZOID	=~	SZDO	1	6
PARI	~~	PARI	0	7
PARC	~~	PARC	0	8
PARO	~~	PARO	0	9
SZTI	~~	SZTI	0	10
SZTC	~~	SZTC	0	11
SZTO	~~	SZTO	0	12
SZDI	~~	SZDI	0	13
SZDC	~~	SZDC	0	14
SZDO	~~	SZDO	0	15
PARANOID	~~	PARANOID	0	16
SCHIZOTYPAL	~~	SCHIZOTYPAL	0	17
SCHIZOID	~~	SCHIZOID	0	18
PARANOID	~~	SCHIZOTYPAL	0	19
PARANOID	~~	SCHIZOID	0	20
SCHIZOTYPAL	~~	SCHIZOID	0	21

Assumptions

Open issue

Model Settings

Outputs

Model Fit Summary

```
lavaan 0.6-7 ended normally after 73 iterations
```

```
  Estimator                                         ML
  Optimization method                           NLMINB
  Number of free parameters                         21

  Number of observations                           250

Model Test User Model:

  Test statistic                               227.660
  Degrees of freedom                                24
  P-value (Chi-square)                           0.000

Model Test Baseline Model:

  Test statistic                              1251.828
  Degrees of freedom                                36
  P-value                                        0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.832
  Tucker-Lewis Index (TLI)                       0.749

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -5047.979
  Loglikelihood unrestricted model (H1)      -4934.149

  Akaike (AIC)                               10137.957
  Bayesian (BIC)                             10211.908
  Sample-size adjusted Bayesian (BIC)        10145.336

Root Mean Square Error of Approximation:

  RMSEA                                          0.184
  90 Percent confidence interval - lower         0.163
  90 Percent confidence interval - upper         0.206
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.063

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured
```

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  PARANOID =~
    PARI              1.000
    PARC              0.942    0.088   10.668    0.000
    PARO              0.711    0.067   10.665    0.000
  SCHIZOTYPAL =~
    SZTI              1.000
    SZTC              0.799    0.071   11.211    0.000
    SZTO              0.725    0.061   11.905    0.000
  SCHIZOID =~
    SZDI              1.000
    SZDC              0.865    0.062   13.979    0.000
    SZDO              0.635    0.045   14.163    0.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  PARANOID ~~
    SCHIZOTYPAL       2.965    0.628    4.722    0.000
    SCHIZOID          2.736    0.599    4.568    0.000
  SCHIZOTYPAL ~~
    SCHIZOID          2.854    0.645    4.425    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .PARI              6.545    0.734    8.920    0.000
   .PARC              2.897    0.454    6.388    0.000
   .PARO              1.660    0.259    6.417    0.000
   .SZTI              5.358    0.691    7.757    0.000
   .SZTC              4.047    0.484    8.360    0.000
   .SZTO              1.641    0.295    5.566    0.000
   .SZDI              4.944    0.568    8.711    0.000
   .SZDC              2.183    0.325    6.713    0.000
   .SZDO              0.967    0.164    5.897    0.000
    PARANOID          6.435    1.096    5.873    0.000
    SCHIZOTYPAL       7.984    1.201    6.649    0.000
    SCHIZOID          7.892    1.118    7.061    0.000
```

Completely Standardized Parameter Estimates

```
Latent Variables:
                    est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  PARANOID =~
    PARI              0.704    0.040   17.590    0.000    0.626    0.783
    PARC              0.815    0.034   23.940    0.000    0.748    0.881
    PARO              0.814    0.034   23.883    0.000    0.747    0.880
```

```
  SCHIZOTYPAL =~
    SZTI              0.774    0.035   22.096    0.000    0.705    0.842
    SZTC              0.747    0.037   20.410    0.000    0.675    0.818
    SZTO              0.848    0.031   27.016    0.000    0.786    0.909
  SCHIZOID =~
    SZDI              0.784    0.030   26.053    0.000    0.725    0.843
    SZDC              0.854    0.025   33.640    0.000    0.805    0.904
    SZDO              0.876    0.024   36.162    0.000    0.828    0.923

Covariances:
                    est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  PARANOID ~~
    SCHIZOTYPAL       0.414    0.065    6.347    0.000    0.286    0.541
    SCHIZOID          0.384    0.065    5.933    0.000    0.257    0.511
  SCHIZOTYPAL ~~
    SCHIZOID          0.360    0.065    5.500    0.000    0.231    0.488

Variances:
                    est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .PARI              0.504    0.056    8.944    0.000    0.394    0.615
   .PARC              0.337    0.055    6.072    0.000    0.228    0.445
   .PARO              0.338    0.055    6.100    0.000    0.229    0.447
   .SZTI              0.402    0.054    7.414    0.000    0.295    0.508
   .SZTC              0.443    0.055    8.104    0.000    0.336    0.550
   .SZTO              0.281    0.053    5.287    0.000    0.177    0.386
   .SZDI              0.385    0.047    8.161    0.000    0.293    0.478
   .SZDC              0.270    0.043    6.217    0.000    0.185    0.355
   .SZDO              0.233    0.042    5.503    0.000    0.150    0.316
    PARANOID          1.000                               1.000    1.000
    SCHIZOTYPAL       1.000                               1.000    1.000
    SCHIZOID          1.000                               1.000    1.000
```

Communality

Table 4: Communality

Variable	Communality
PARC	0.66
PARI	0.50
PARO	0.66
SZDC	0.73
SZDI	0.61
SZDO	0.77
SZTC	0.56
SZTI	0.60
SZTO	0.72

Factor Discriminant Validity

Table 5: Factor Discriminant Validity Test at Cutoff 0.85

			Factor Correlation	Chisq diff	Df diff	P-Value
PARANOID	~~	SCHIZOTYPAL	0.414	92.792	1	<0.001
PARANOID	~~	SCHIZOID	0.384	110.221	1	<0.001
SCHIZOTYPAL	~~	SCHIZOID	0.360	120.010	1	<0.001

Factor Reliability

Table 6: Factor Reliability

	PARANOID	SCHIZOTYPAL	SCHIZOID	total
Omega (Bentler)	0.80	0.82	0.86	0.89
Omega (McDonald)	0.80	0.82	0.86	0.89
AVE	0.58	0.61	0.68	0.62

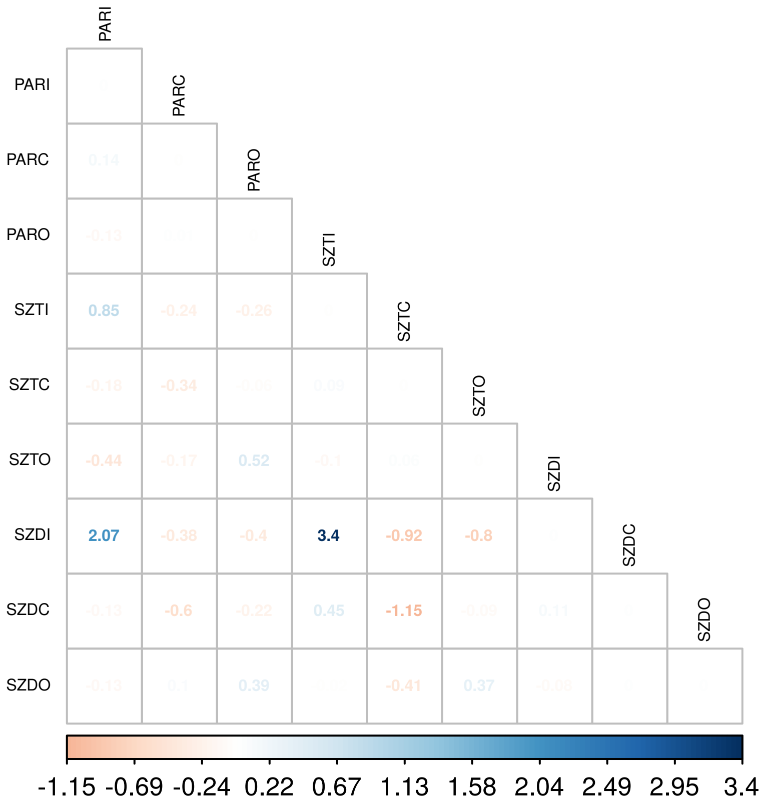
Observed Covariance Matrix

	PARI	PARC	PARO	SZTI	SZTC	SZTO	SZDI	SZDC	SZDO
PARI	12.98								
PARC	6.21	8.61							
PARO	4.45	4.32	4.91						
SZTI	3.82	2.55	1.85	13.34					
SZTC	2.19	1.89	1.62	6.47	9.14				
SZTO	1.71	1.85	2.05	5.68	4.68	5.83			
SZDI	4.8	2.2	1.55	6.26	1.37	1.26	12.84		
SZDC	2.23	1.63	1.46	2.92	0.83	1.7	6.94	8.09	
SZDO	1.61	1.73	1.63	1.79	1.03	1.68	4.93	4.34	4.14

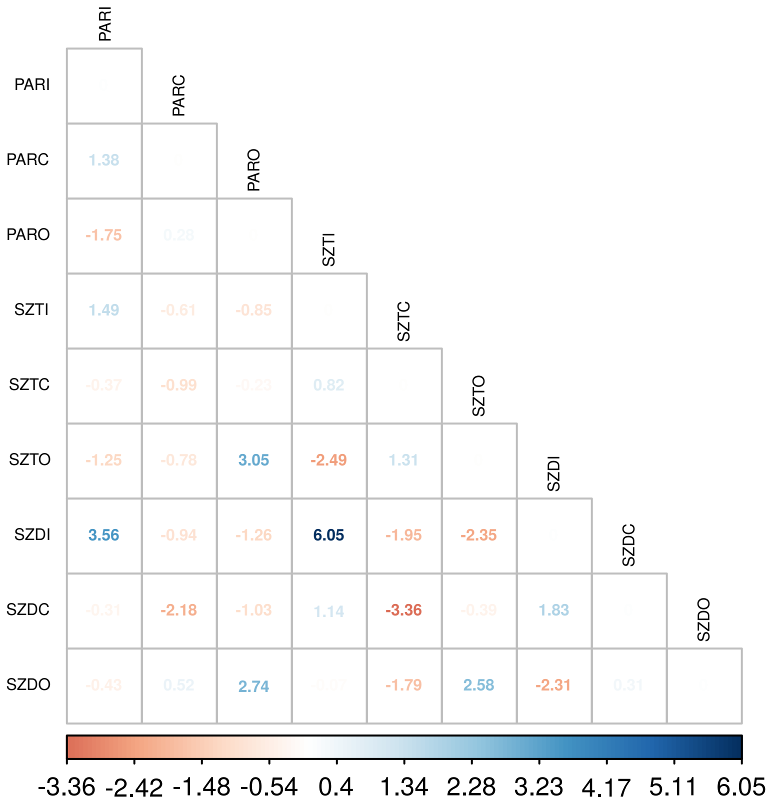
Model-Implied Covariance Matrix

	PARI	PARC	PARO	SZTI	SZTC	SZTO	SZDI	SZDC	SZDO
PARI	12.98								
PARC	6.06	8.61							
PARO	4.57	4.31	4.91						
SZTI	2.96	2.79	2.11	13.34					
SZTC	2.37	2.23	1.68	6.38	9.14				
SZTO	2.15	2.02	1.53	5.79	4.62	5.83			
SZDI	2.74	2.58	1.94	2.85	2.28	2.07	12.84		
SZDC	2.37	2.23	1.68	2.47	1.97	1.79	6.83	8.09	
SZDO	1.74	1.64	1.23	1.81	1.45	1.31	5.01	4.33	4.14

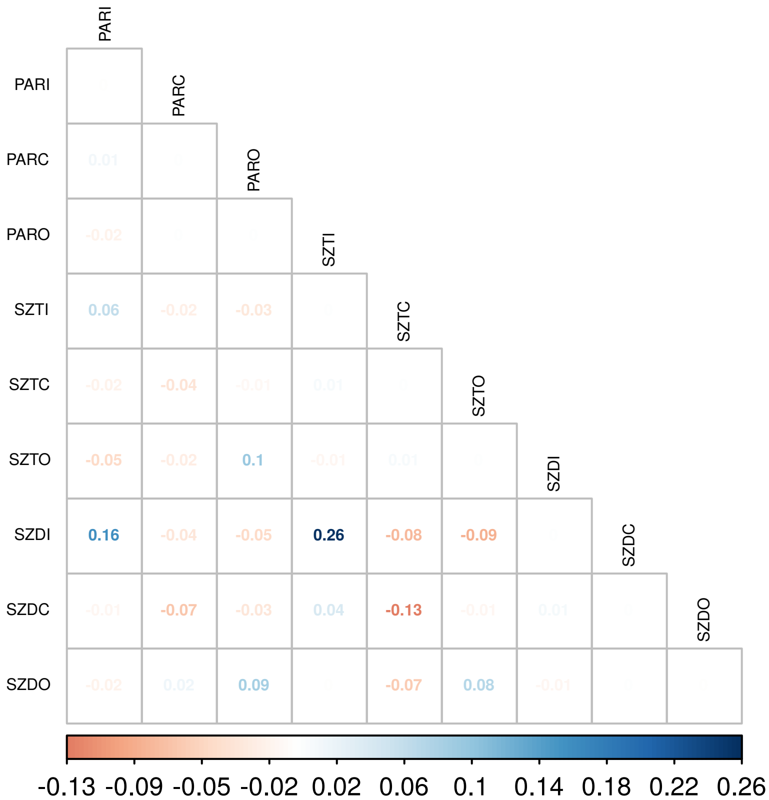
Residual Covariance Matrix



Standardized Residual Matrix



Residual Correlation Matrix



Modification Indices

Table 7: Modification Indices With Respect To Error Covariances

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
SZTI	~~	SZDI	94.979	3.998	0.1	0.057	**(m)**
SZTO	~~	SZDI	54.863	-1.897	0.1	0.068	**(m)**
SZTI	~~	SZDO	37.616	-1.276	0.1	0.077	**(m)**
PARI	~~	SZDI	35.991	2.606	0.1	0.056	**(m)**
SZTO	~~	SZDO	28.936	0.701	0.1	0.120	**(m)**
PARO	~~	SZTO	19.962	0.735	0.1	0.093	**(m)**
PARI	~~	SZDO	17.437	-0.919	0.1	0.074	**(m)**
PARO	~~	SZDO	14.975	0.479	0.1	0.128	**(m)**
PARO	~~	SZDI	14.104	-0.913	0.1	0.070	**(m)**
PARO	~~	SZTI	11.213	-0.877	0.1	0.067	**(m)**

Note:
Maximum 10 modification indices in descending order of their magnitude are listed.

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: 45 - 21 = 24 . The 24 degrees of freedom indicate an over-identified model, fact which basically enables further analysis and interpretation.

The test statistic with the value 227.66 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 <0.001 suggesting that the model may not be acceptable for the data. The Chi-square model fit index is based on a very stringent statistical hypothesis which may have no practical relevance. We will consider it only in connection with other model fit indices.

Model Test Baseline Model
The test statistic with the value 1251.828 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 0.184. The upper bound of the 90% confidence interval of the RMSEA is 0.206 and greater or equal than the threshold value 0.1, suggesting a poor 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 0.06 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 0.832 which is smaller or equal than the threshold value 0.90, suggesting a poor 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 0.749 which is smaller or equal than the threshold value 0.90, suggesting a poor model fit.

Summary of the Goodness of Fit Indices

At least the comparative fit indices CFI and TLI suggest a poor model fit. Therefore, we assume a poor model fit and proceed by diagnosing the sources of possible misspecification.

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.
Following variable pairs have 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]. In these cases, the covariance relationship between the involved variables is probably underestimated:

Table 8: Pair(s) with Underestimated Covariance

Pair 1	PARI	SZDI
Pair 2	PARO	SZDO
Pair 3	SZDO	SZTO
Pair 4	SZTI	SZDI
Pair 5	SZTO	SZDO

Following variable pairs have 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]. In these cases, the covariance relationship between the involved variables is probably overestimated:

Table 9: Pair(s) with Overestimated Covariance

Pair 1	SZDC	SZTC
Pair 2	SZTC	SZDC

Depending on the sample size, the misspecification detected by the analysis of the residual covariance resp. correlation matrices can be statistically significant but not relevant and in practice negligible. This is matter of subject in the next section(s).

Modification Indices

In the interpretation of the modification indices table(s) 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 error or factor correlations, respectively 0.4 for factor loadings. These are characterized in the decision column by the label "epc:m". For more information with regard to the labels of the decision column, please consult the Appendix.
We remark that these conditions are not fulfilled for modification indices with respect to error covariances. Therefore, there exist no significant and relevant modification indices with respect to error covariances.
We remark that there exist no modification indices with respect to factor loadings.
We remark that there exist no modification indices with respect to factor covariances.

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 10: Check Completely Standardized Factor Loadings

Latent Variable	Observed Variable	Loading[1]	P-Value	Significant?[2]	Relevance[3]	Sign[4]	Check
PARANOID	PARI	0.70	<0.001	Yes	**	—	Ok
PARANOID	PARC	0.81	<0.001	Yes	***	—	Ok
PARANOID	PARO	0.81	<0.001	Yes	***	—	Ok
SCHIZOTYPAL	SZTI	0.77	<0.001	Yes	**	—	Ok
SCHIZOTYPAL	SZTC	0.75	<0.001	Yes	**	—	Ok
SCHIZOTYPAL	SZTO	0.85	<0.001	Yes	***	—	Ok
SCHIZOID	SZDI	0.78	<0.001	Yes	**	—	Ok
SCHIZOID	SZDC	0.85	<0.001	Yes	***	—	Ok
SCHIZOID	SZDO	0.88	<0.001	Yes	***	—	Ok

[1] The completely standardized factor loading can be interpreted as the correlation with the factor.
[2] Completely standardized factor loading significance at 5% type I error.
[3] Stars correspond to factor loadings cutoff-values: 0.4, 0.6, 0.8.
[4] No (correct) information available. We assume the signs of the factor loadings correspond to your expectation.
[5] _____
[6] _____
[7] _____

Moreover, we remark that the significance test results for the completely standardized factor loadings from above coincide to those of the unstandardized factor loadings (within section "Model Fit Summary", for non-marker variables).

We proceed by interpreting the (unstandardized) factor loadings from the "Model Fit Summary" section:

Table 11: Interpretation of Unstandardized Factor Loadings

Interpretation of Unstandardized Factor Loadings
A 1-unit increase in PARANOID leads to a 1.00 -unit increase in the PARI

Table 11: Interpretation of Unstandardized Factor Loadings (continued)

Interpretation of Unstandardized Factor Loadings
A 1-unit increase in PARANOID leads to a 0.94 -unit increase in the PARC
A 1-unit increase in PARANOID leads to a 0.71 -unit increase in the PARO
A 1-unit increase in SCHIZOTYPAL leads to a 1.00 -unit increase in the SZTI
A 1-unit increase in SCHIZOTYPAL leads to a 0.80 -unit increase in the SZTC
A 1-unit increase in SCHIZOTYPAL leads to a 0.72 -unit increase in the SZTO
A 1-unit increase in SCHIZOID leads to a 1.00 -unit increase in the SZDI
A 1-unit increase in SCHIZOID leads to a 0.87 -unit increase in the SZDC
A 1-unit increase in SCHIZOID leads to a 0.63 -unit increase in the SZDO

Factor Discriminant Validity

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 0.41 which we regard as a proof of a reasonable discriminant validity. Moreover, the statistical test(s) for factor discriminant validity are statistically significant at 5% type I error.

Error Variances

We summarize the interpretation of the error variances and communalities in the next table:

Table 12: Completely Standardized Error Variances and Communality

Observed Variable	Error Variance[1]	Communality[2,3]	P-Value	Significant Error Variance?[4]
PARC	0.34	0.66	<0.001	Yes
PARI	0.50	0.50	<0.001	Yes
PARO	0.34	0.66	<0.001	Yes
SZDC	0.27	0.73	<0.001	Yes
SZDI	0.39	0.61	<0.001	Yes
SZDO	0.23	0.77	<0.001	Yes
SZTC	0.44	0.56	<0.001	Yes
SZTI	0.40	0.60	<0.001	Yes
SZTO	0.28	0.72	<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.

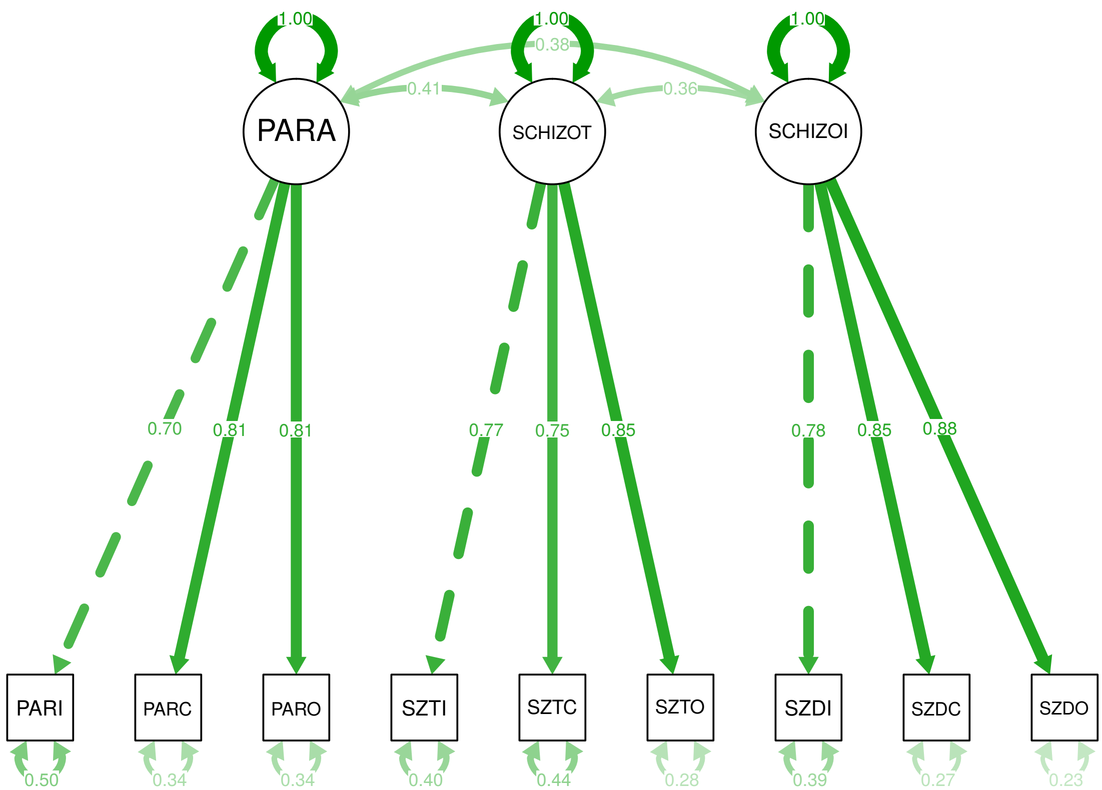
Factor Reliability

The table "Factor Reliability" contains the omega measures of factor reliability given by Bentler (Bentler, 1972, 2009) and McDonald (McDonald, 1999) and the average variance extracted (AVE). The interpretatibility of the reliability measures depend on the specific research context. In some fields of research, omega values greater or equal than 0.6 and AVE values greater or equal than 0.5 (fulfilled by and large in your case) could be sufficient.

Final Summary

In our final evaluation, we distinguish between following model quality categories: acceptable, non-acceptable or uncertain.
Considering the goodness-of-fit indices, the model is non-acceptable. Please reconsider your data and the theory behind.

Path Diagram



APPENDIX

Decision Column of the Modification Indices Table

```
not mi.significant & not high.power := "(i)"
mi.significant & not high.power := "**(m)**"
not mi.significant & high.power := "(nm)"
mi.significant & high.power & not epc.high := "epc:nm"
mi.significant & high.power & epc.high := "*epc:m*"
```

References