

Confirmatory Factor Analysis

Statsomat.com

08 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
case7.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
hm
nr
wo
gc
tr
sm
ma
ps

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
Sequent	=~	hm	1	0
Sequent	=~	nr	1	1
Sequent	=~	wo	1	2
Simultan	=~	gc	1	0
Simultan	=~	tr	1	3
Simultan	=~	sm	1	4
Simultan	=~	ma	1	5
Simultan	=~	ps	1	6
hm	~~	hm	0	7
nr	~~	nr	0	8
wo	~~	wo	0	9
gc	~~	gc	0	10
tr	~~	tr	0	11
sm	~~	sm	0	12
ma	~~	ma	0	13
ps	~~	ps	0	14
Sequent	~~	Sequent	0	15
Simultan	~~	Simultan	0	16
Sequent	~~	Simultan	0	17

Assumptions

Open issue

Model Settings

Outputs

Model Fit Summary

lavaan 0.6-7 ended normally after 43 iterations

Estimator	ML
Optimization method	NLMINB
Number of free parameters	17
Number of observations	250

Model Test User Model:

Test statistic	47.906
Degrees of freedom	19
P-value (Chi-square)	0.000

Model Test Baseline Model:

Test statistic	622.920
Degrees of freedom	28
P-value	0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)	0.951
Tucker-Lewis Index (TLI)	0.928

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-4724.805
Loglikelihood unrestricted model (H1)	-4700.852
Akaike (AIC)	9483.611
Bayesian (BIC)	9543.476
Sample-size adjusted Bayesian (BIC)	9489.584

Root Mean Square Error of Approximation:

RMSEA	0.078
90 Percent confidence interval - lower	0.051
90 Percent confidence interval - upper	0.106
P-value RMSEA <= 0.05	0.046

Standardized Root Mean Square Residual:

SRMR	0.072
------	-------

Parameter Estimates:

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

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
Sequent =~				
hm	1.000			

nr	1.147	0.162	7.089	0.000
wo	1.388	0.196	7.088	0.000
Simultan =~				
gc	1.000			
tr	1.445	0.203	7.102	0.000
sm	2.029	0.299	6.777	0.000
ma	1.212	0.190	6.391	0.000
ps	1.727	0.237	7.291	0.000

Covariances:

	Estimate	Std.Err	z-value	P(> z)
Sequent ~~				
Simultan	1.272	0.290	4.380	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.hm	8.673	0.840	10.327	0.000
.nr	2.000	0.370	5.401	0.000
.wo	2.905	0.541	5.368	0.000
.gc	5.425	0.524	10.354	0.000
.tr	3.429	0.410	8.362	0.000
.sm	10.008	1.076	9.302	0.000
.ma	5.110	0.517	9.881	0.000
.ps	3.486	0.481	7.247	0.000
Sequent	2.841	0.750	3.789	0.000
Simultan	1.836	0.475	3.868	0.000

Completely Standardized Parameter Estimates

Latent Variables:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
Sequent =~						
hm	0.497	0.055	8.979	0.000	0.388	0.605
nr	0.807	0.041	19.501	0.000	0.726	0.888
wo	0.808	0.041	19.541	0.000	0.727	0.889
Simultan =~						
gc	0.503	0.054	9.236	0.000	0.396	0.610
tr	0.726	0.039	18.405	0.000	0.649	0.804
sm	0.656	0.044	14.788	0.000	0.569	0.743
ma	0.588	0.049	11.981	0.000	0.492	0.684
ps	0.782	0.036	21.783	0.000	0.711	0.852

Covariances:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
Sequent ~~						
Simultan	0.557	0.060	9.331	0.000	0.440	0.674

Variances:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
.hm	0.753	0.055	13.705	0.000	0.646	0.861
.nr	0.349	0.067	5.220	0.000	0.218	0.480
.wo	0.347	0.067	5.188	0.000	0.216	0.478
.gc	0.747	0.055	13.641	0.000	0.640	0.854
.tr	0.472	0.057	8.235	0.000	0.360	0.585
.sm	0.570	0.058	9.785	0.000	0.456	0.684
.ma	0.654	0.058	11.343	0.000	0.541	0.767
.ps	0.389	0.056	6.931	0.000	0.279	0.499
Sequent	1.000				1.000	1.000
Simultan	1.000				1.000	1.000

Communality

Table 4: Communality

Variable	Communality
gc	0.25
hm	0.25
ma	0.35
nr	0.65
ps	0.61
sm	0.43
tr	0.53
wo	0.65

Factor Discriminant Validity

Table 5: Factor Discriminant Validity Test at Cutoff 0.85

			Factor Correlation	Chisq diff	Df diff	P-Value
Sequent	~~	Simultan	0.557	35.746	1	<0.001

Factor Reliability

Table 6: Factor Reliability

	Sequent	Simultan	total
Omega (Bentler)	0.72	0.79	0.83
Omega (McDonald)	0.74	0.79	0.79
AVE	0.47	0.44	0.45

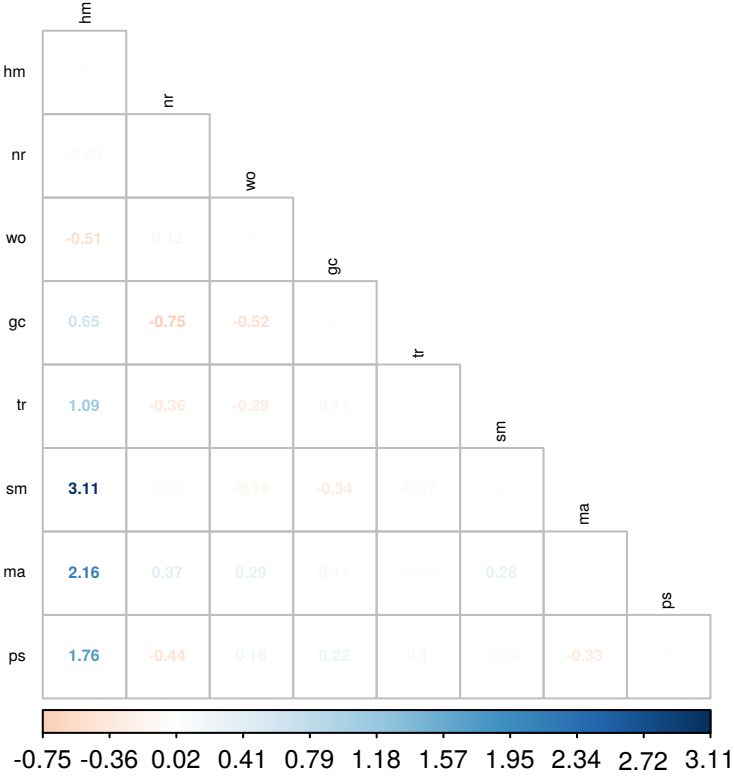
Observed Covariance Matrix

hm	11.51						
nr	3.17	5.74					
wo	3.44	4.64	8.38				
gc	1.92	0.71	1.25	7.26			
tr	2.93	1.74	2.26	2.76	7.26		
sm	5.69	2.91	3.4	3.39	5.31	17.57	
ma	3.7	2.14	2.43	2.33	3.16	4.8	7.81
ps	3.96	2.08	3.21	3.39	4.68	6.4	3.51
							8.96

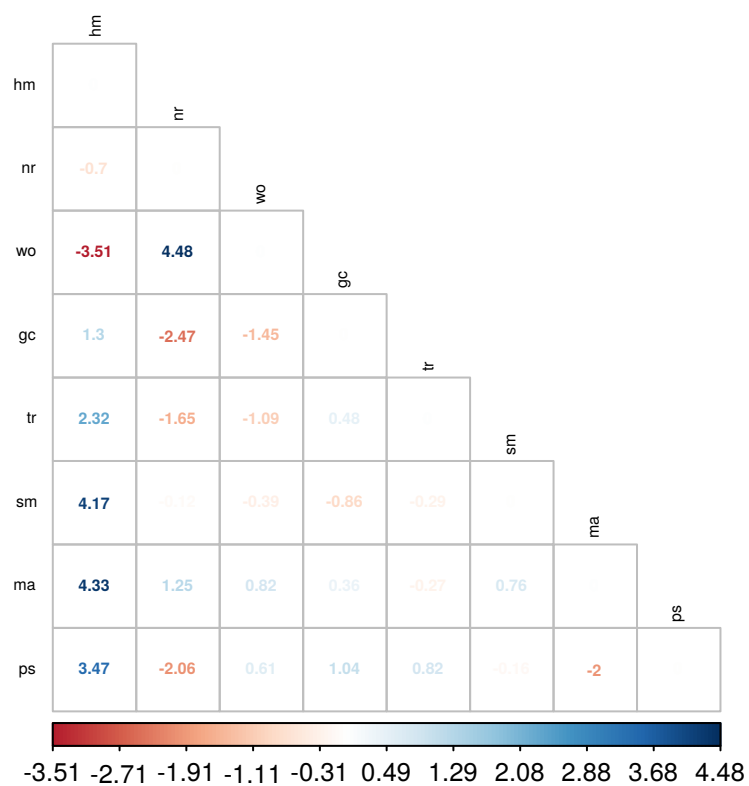
Model-Implied Covariance Matrix

hm	11.51							
nr	3.26	5.74						
wo	3.94	4.52	8.38					
gc	1.27	1.46	1.77	7.26				
tr	1.84	2.11	2.55	2.65	7.26			
sm	2.58	2.96	3.58	3.73	5.38	17.57		
ma	1.54	1.77	2.14	2.23	3.22	4.52	7.81	
ps	2.2	2.52	3.05	3.17	4.58	6.44	3.85	8.96

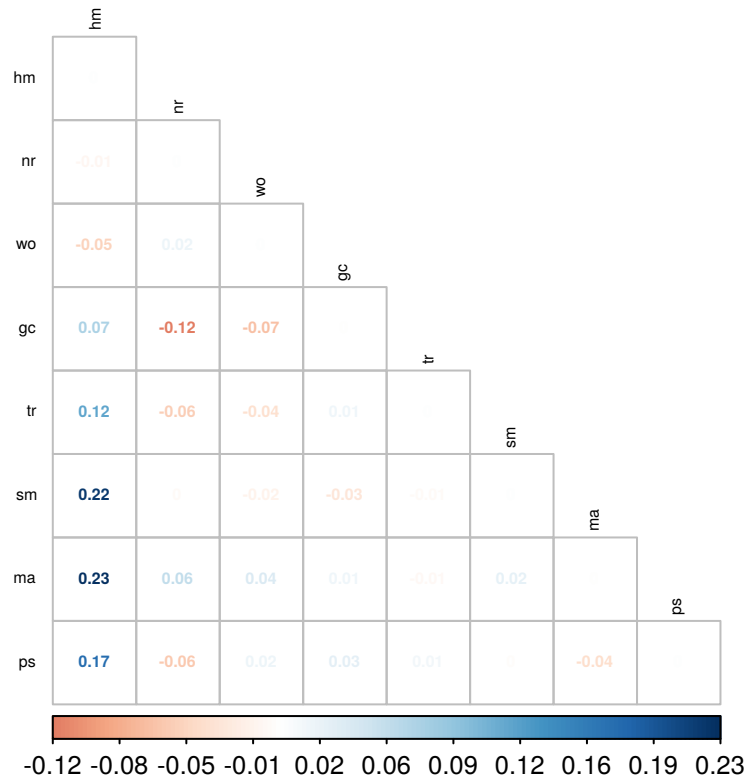
Residual Covariance Matrix



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
nr	~~	wo	25.121	4.745	0.1	0.051	** (m) **
hm	~~	wo	8.766	-1.748	0.1	0.053	** (m) **
hm	~~	sm	6.059	1.611	0.1	0.053	** (m) **
hm	~~	ma	4.748	0.996	0.1	0.056	** (m) **
nr	~~	ps	3.934	-0.503	0.1	0.068	** (m) **
ma	~~	ps	3.408	-0.733	0.1	0.057	(i)

Note:

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

Table 8: Modification Indices With Respect To Factor Loadings

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
Simultan	=~	hm	25.121	1.054	0.4	0.477	** (m) **
Simultan	=~	nr	8.766	-0.510	0.4	0.641	** (m) **
Sequent	=~	ma	4.058	0.269	0.4	0.849	epc: nm
Sequent	=~	gc	3.627	-0.254	0.4	0.849	(nm)

Note:

Table 8: Modification Indices With Respect To Factor Loadings (continued)

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
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: $36 - 17 = 19$. The 19 degrees of freedom indicate an over-identified model, fact which basically enables further analysis and interpretation.

The test statistic with the value 47.906 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 622.92 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 0.078. The upper bound of the 90% confidence interval of the RMSEA is 0.106 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 0.07 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 0.951 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 0.928 which is greater or equal than the threshold value 0.90, suggesting an acceptable model fit.

Summary of the Goodness of Fit Indices

The TLI model fit index suggests an acceptable model fit. The SRMR model fit index suggests an acceptable model fit. The Chi-square model fit index and the RMSEA suggest a poor model fit. The goodness-of-fit of the model is uncertain. We 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 9: Pair(s) with Underestimated Covariance

Pair 1	hm	ma
Pair 2	hm	ps
Pair 3	hm	sm
Pair 4	hm	tr
Pair 5	ma	ma
Pair 6	nr	wo

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 10: Pair(s) with Overestimated Covariance

Pair 1	hm	wo
Pair 2	gc	nr

Depending on the sample size, the misspecification detected by the analysis of the residual covariance resp. corre-

lation 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 χ^2 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 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.

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 11: Check Completely Standardized Factor Loadings

Latent Variable	Observed Variable	Loading ¹	P-Value	Significant? ²	Relevance ³	Sign ⁴	Check
Sequent	hm	0.50	<0.001	Yes	*	—	Ok
Sequent	nr	0.81	<0.001	Yes	***	—	Ok
Sequent	wo	0.81	<0.001	Yes	***	—	Ok
Simultan	gc	0.50	<0.001	Yes	*	—	Ok
Simultan	tr	0.73	<0.001	Yes	**	—	Ok
Simultan	sm	0.66	<0.001	Yes	**	—	Ok
Simultan	ma	0.59	<0.001	Yes	*	—	Ok
Simultan	ps	0.78	<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 _____

Table 11: Check Completely Standardized Factor Loadings (continued)

Latent Variable	Observed Variable	Loading ¹	P-Value	Significant? ²	Relevance ³	Sign ⁴	Check
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 12: Interpretation of Unstandardized Factor Loadings

Interpretation of Unstandardized Factor Loadings
A 1-unit increase in Sequent leads to a 1.00 -unit increase in the hm
A 1-unit increase in Sequent leads to a 1.15 -unit increase in the nr
A 1-unit increase in Sequent leads to a 1.39 -unit increase in the wo
A 1-unit increase in Simultan leads to a 1.00 -unit increase in the gc
A 1-unit increase in Simultan leads to a 1.44 -unit increase in the tr
A 1-unit increase in Simultan leads to a 2.03 -unit increase in the sm
A 1-unit increase in Simultan leads to a 1.21 -unit increase in the ma
A 1-unit increase in Simultan leads to a 1.73 -unit increase in the ps

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 (absolute) estimated factor intercorrelation within the section “Completely Standardized Parameter Estimates” is 0.56 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 13: Completely Standardized Error Variances and Commuality

Observed Variable	Error Variance ¹	Communality ²³	P-Value	Significant Error Variance? ⁴
gc	0.75	0.25	<0.001	Yes
hm	0.75	0.25	<0.001	Yes
ma	0.65	0.35	<0.001	Yes
nr	0.35	0.65	<0.001	Yes
ps	0.39	0.61	<0.001	Yes
sm	0.57	0.43	<0.001	Yes
tr	0.47	0.53	<0.001	Yes
wo	0.35	0.65	<0.001	Yes

¹ Can be interpreted as proportion of unexplained variance by the latent factor(s) (%).

² Corresponds to the squared factor loading.

Table 13: Completely Standardized Error Variances and Communality
(continued)

Observed Variable	Error Variance ¹	Communality ²³	P-Value	Significant Error Variance ²⁴
-------------------	-----------------------------	---------------------------	---------	--

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

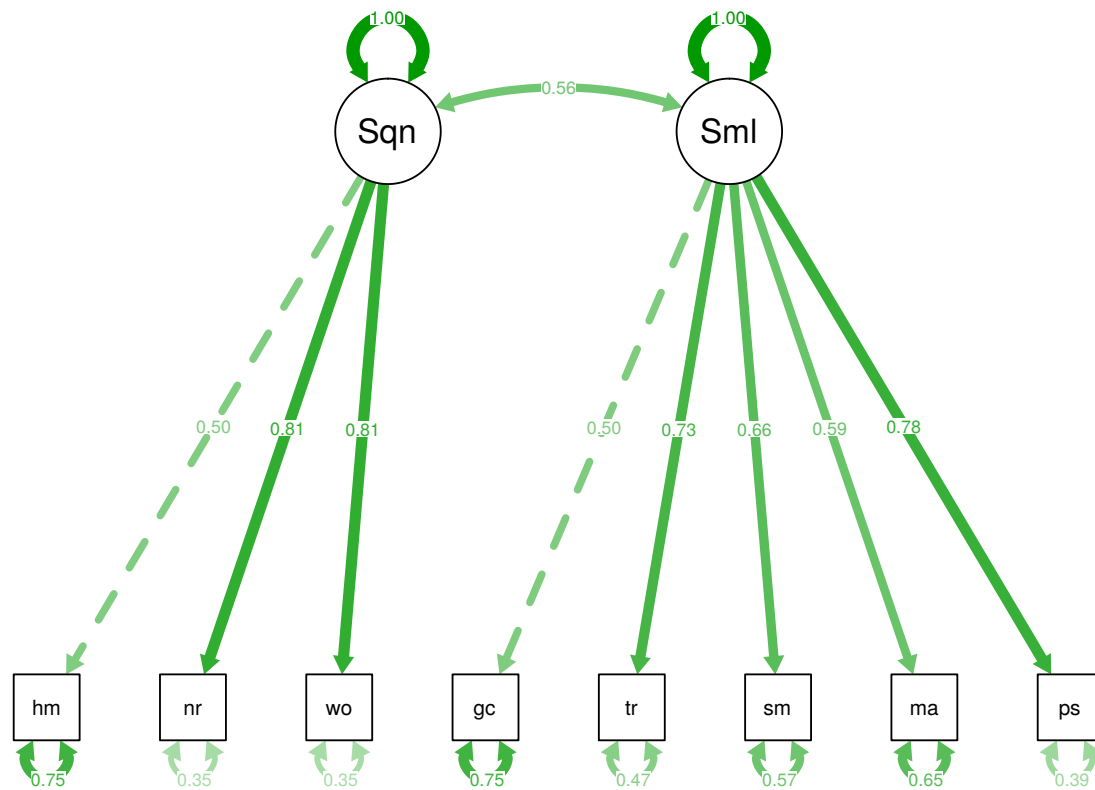
The table “Factor Reliability” contains the omega measures of factor reliability given by Bentler (Bentler, 1972, 2009) and McDonald (McDonald, 1999) and the average variance extracted (AVE). The interpretability of the reliability measures depend on the specific research context. Nevertheless, omega values below 0.6 or AVE values below 0.5 (at least one of these existent in your case) should be regarded with criticism. The factor reliability estimates are not further considered in the final summary.

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 quality of the model is uncertain. Since we cannot identify (significant and relevant) localized areas of ill fit, we assume that the model is acceptable but strongly recommend to replicate the CFA in an independent sample to validate the results. Please consider also that this summary depends on (some) hard-coded cutoff-values which may be too liberal or too conservative for your research area.

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