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
case1.csv

Your selection for the encoding: Auto
Your selection for the decimal character: Auto
Observations (rows with at least one non-missing value): 301
Variables (columns with at least one non-missing value): 9
Variables considered continuous: 8

Variables considered continuous
x1
x2
x3
x5
x6
x7
x8
x9

Numerical variables considered binary or ordinal: 1

Numerical variables considered binary or ordinal
x4

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
visual	=~	x1	1	0
visual	=~	x2	1	1
visual	=~	x3	1	2
textual	=~	x4	1	0
textual	=~	x5	1	3
textual	=~	x6	1	4
speed	=~	x7	1	0
speed	=~	x8	1	5
speed	=~	x9	1	6
x1	~~	x1	0	7
x2	~~	x2	0	8
x3	~~	x3	0	9
x4	~~	x4	0	10
x5	~~	x5	0	11
x6	~~	x6	0	12
x7	~~	x7	0	13
x8	~~	x8	0	14
x9	~~	x9	0	15
visual	~~	visual	0	16
textual	~~	textual	0	17
speed	~~	speed	0	18
visual	~~	textual	0	19
visual	~~	speed	0	20
textual	~~	speed	0	21
x1	~1		0	22
x2	~1		0	23
x3	~1		0	24
x4	~1		0	25
x5	~1		0	26
x6	~1		0	27
x7	~1		0	28
x8	~1		0	29
x9	~1		0	30
visual	~1		0	0
textual	~1		0	0
speed	~1		0	0

Assumptions

Open issue

Model Settings

Outputs

Model Fit Summary

```
lavaan 0.6-7 ended normally after 35 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of free parameters                         30

  Number of observations                           301
  Number of missing patterns                         1

Model Test User Model:

  Test statistic                                85.306
  Degrees of freedom                                24
  P-value (Chi-square)                           0.000

Model Test Baseline Model:

  Test statistic                               918.852
  Degrees of freedom                                36
  P-value                                        0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.931
  Tucker-Lewis Index (TLI)                       0.896

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -3737.745
  Loglikelihood unrestricted model (H1)      -3695.092

  Akaike (AIC)                                7535.490
  Bayesian (BIC)                              7646.703
  Sample-size adjusted Bayesian (BIC)         7551.560
```

```
Root Mean Square Error of Approximation:

  RMSEA                                          0.092
  90 Percent confidence interval - lower         0.071
  90 Percent confidence interval - upper         0.114
  P-value RMSEA <= 0.05                          0.001

Standardized Root Mean Square Residual:

  SRMR                                           0.060

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Observed
  Observed information based on                Hessian

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  visual =~
    x1                1.000
    x2                0.554    0.109    5.066    0.000
    x3                0.729    0.117    6.220    0.000
  textual =~
    x4                1.000
    x5                1.113    0.065   17.128    0.000
    x6                0.926    0.056   16.481    0.000
  speed =~
    x7                1.000
    x8                1.180    0.150    7.851    0.000
    x9                1.082    0.195    5.543    0.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  visual ~~
    textual           0.408    0.080    5.124    0.000
    speed             0.262    0.055    4.735    0.000
  textual ~~
    speed             0.173    0.049    3.518    0.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .x1                4.936    0.067   73.473    0.000
   .x2                6.088    0.068   89.855    0.000
   .x3                2.250    0.065   34.579    0.000
   .x4                3.061    0.067   45.694    0.000
   .x5                4.341    0.074   58.452    0.000
   .x6                2.186    0.063   34.667    0.000
```

```
   .x7                4.186    0.063   66.766    0.000
   .x8                5.527    0.058   94.854    0.000
   .x9                5.374    0.058   92.546    0.000
    visual            0.000
    textual           0.000
    speed             0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .x1                0.549    0.119    4.612    0.000
   .x2                1.134    0.104   10.875    0.000
   .x3                0.844    0.095    8.881    0.000
   .x4                0.371    0.048    7.739    0.000
   .x5                0.446    0.058    7.703    0.000
   .x6                0.356    0.043    8.200    0.000
   .x7                0.799    0.088    9.130    0.000
   .x8                0.488    0.092    5.321    0.000
   .x9                0.566    0.091    6.250    0.000
    visual            0.809    0.150    5.404    0.000
    textual           0.979    0.112    8.729    0.000
    speed             0.384    0.092    4.168    0.000
```

Completely Standardized Parameter Estimates

```
Latent Variables:
                    est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  visual =~
    x1                0.772    0.058   13.416    0.000    0.659    0.885
    x2                0.424    0.063    6.752    0.000    0.301    0.547
    x3                0.581    0.058    9.942    0.000    0.467    0.696
  textual =~
    x4                0.852    0.023   37.612    0.000    0.807    0.896
    x5                0.855    0.022   38.530    0.000    0.812    0.899
    x6                0.838    0.024   35.598    0.000    0.792    0.884
  speed =~
    x7                0.570    0.058    9.767    0.000    0.455    0.684
    x8                0.723    0.062   11.608    0.000    0.601    0.845
    x9                0.665    0.066   10.063    0.000    0.535    0.795

Covariances:
                    est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  visual ~~
    textual           0.459    0.063    7.225    0.000    0.334    0.583
    speed             0.471    0.086    5.457    0.000    0.302    0.640
  textual ~~
    speed             0.283    0.071    3.959    0.000    0.143    0.423
```

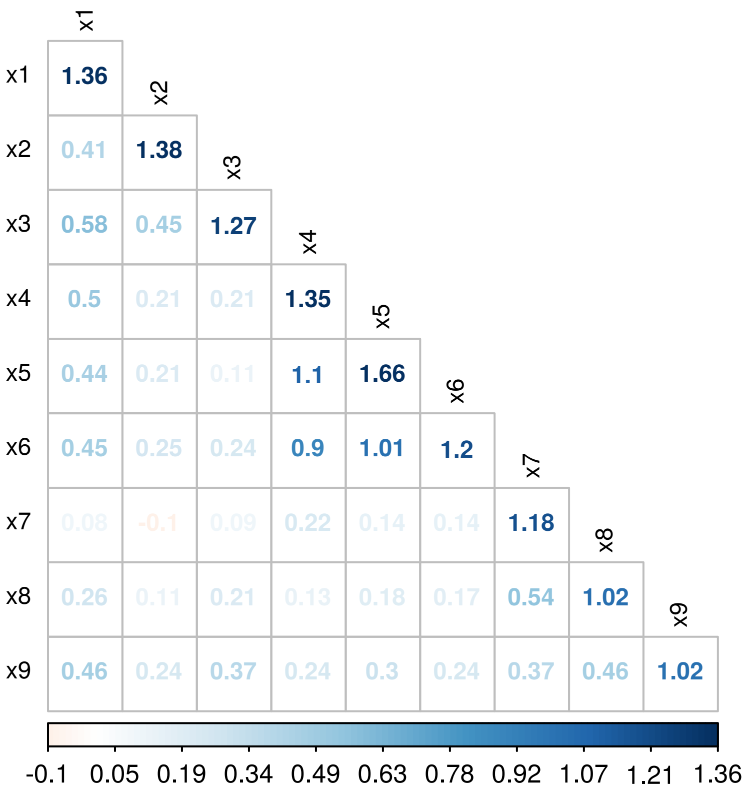
```
Intercepts:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .x1              -0.000    0.058   -0.000    1.000   -0.113    0.113
   .x2              -0.000    0.058   -0.000    1.000   -0.113    0.113
   .x3               0.000    0.058    0.000    1.000   -0.113    0.113
   .x4              -0.000    0.058   -0.000    1.000   -0.113    0.113
   .x5               0.000    0.058    0.000    1.000   -0.113    0.113
   .x6              -0.000    0.058   -0.000    1.000   -0.113    0.113
   .x7              -0.000    0.058   -0.000    1.000   -0.113    0.113
   .x8               0.000    0.058    0.000    1.000   -0.113    0.113
   .x9              -0.000    0.058   -0.000    1.000   -0.113    0.113
    visual           0.000                               0.000    0.000
    textual          0.000                               0.000    0.000
    speed            0.000                               0.000    0.000

Variances:
                   est.std  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .x1               0.404    0.089    4.551    0.000    0.230    0.578
   .x2               0.821    0.053   15.438    0.000    0.716    0.925
   .x3               0.662    0.068    9.748    0.000    0.529    0.795
   .x4               0.275    0.039    7.126    0.000    0.199    0.350
   .x5               0.269    0.038    7.084    0.000    0.194    0.343
   .x6               0.298    0.039    7.546    0.000    0.220    0.375
   .x7               0.676    0.066   10.173    0.000    0.545    0.806
   .x8               0.477    0.090    5.298    0.000    0.301    0.654
   .x9               0.558    0.088    6.346    0.000    0.385    0.730
    visual           1.000                               1.000    1.000
    textual          1.000                               1.000    1.000
    speed            1.000                               1.000    1.000
```

Communality

Table 1: Communality ← Explanation is missing

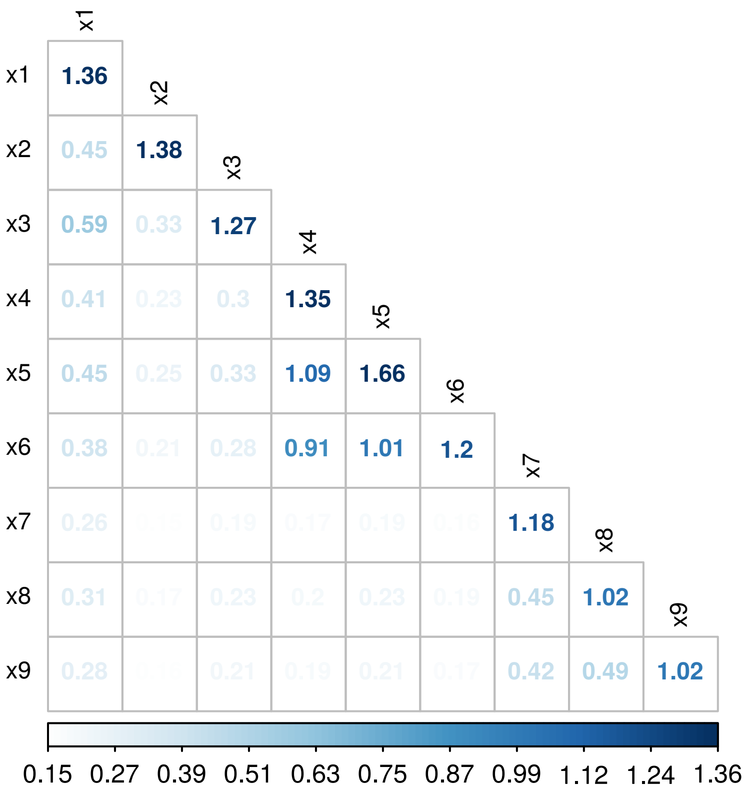
Variable	Communality
x1	0.60
x2	0.18
x3	0.34
x4	0.73
x5	0.73
x6	0.70
x7	0.32
x8	0.52
x9	0.44

Observed Covariance Matrix



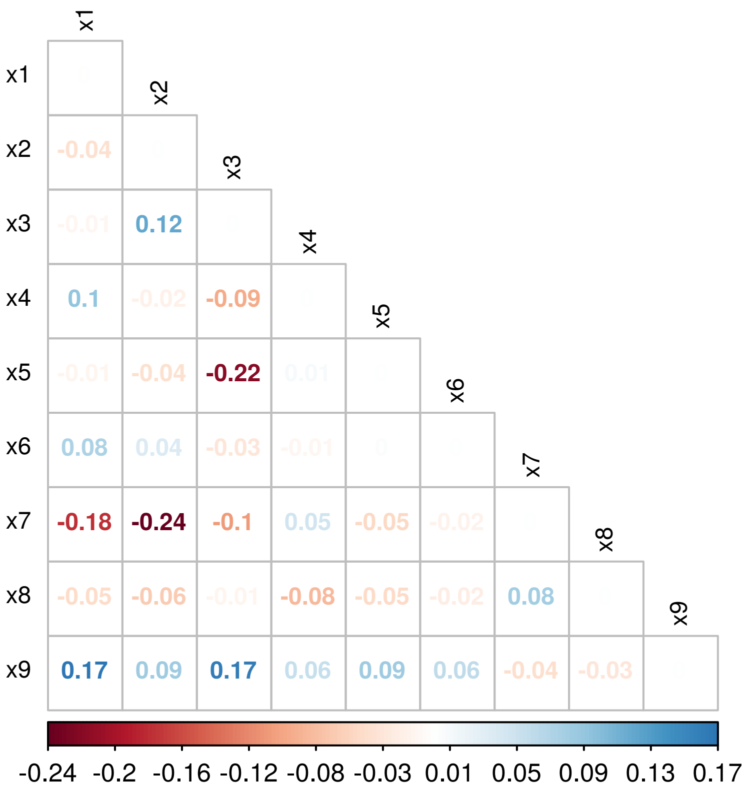
Model

~~Fitted~~ Covariance Matrix

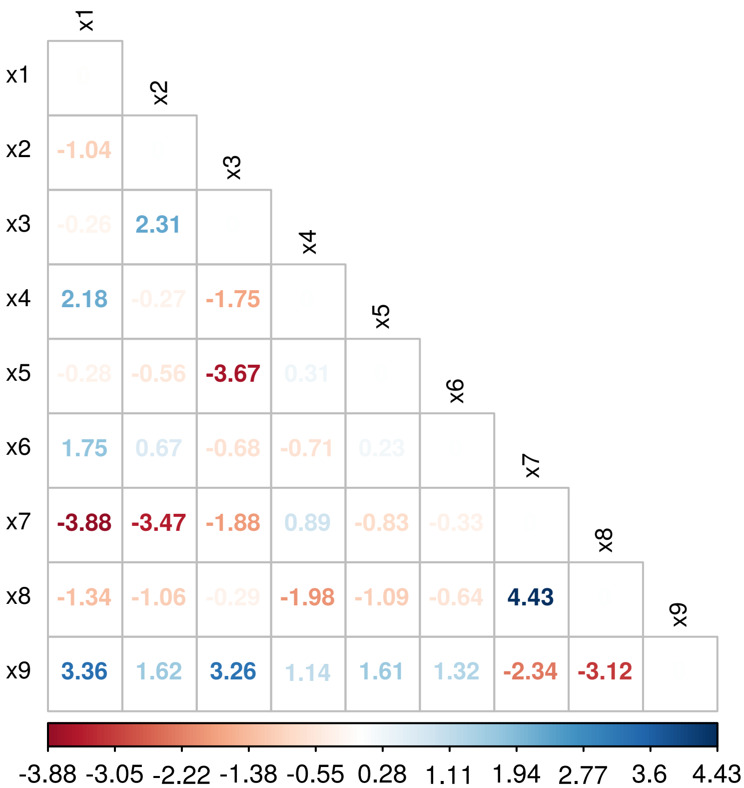


?

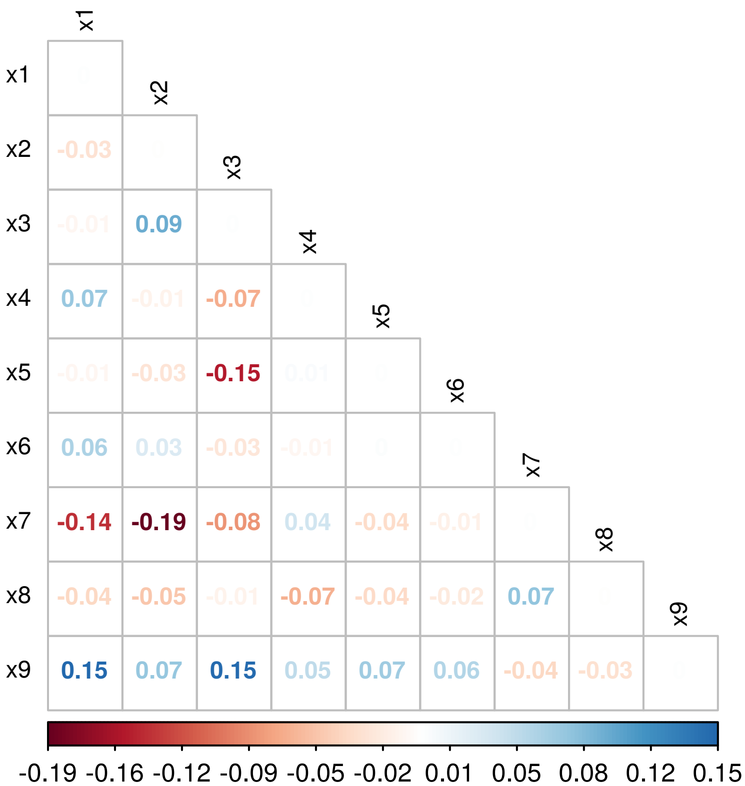
Residual Covariance Matrix



Standardized Residual Matrix



Residual Correlation Matrix



Modification Indices

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

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
x7	~~	x8	34.145	0.536	0.1	0.193	**(m)**
x8	~~	x9	14.946	-0.423	0.1	0.150	**(m)**
x2	~~	x7	8.918	-0.183	0.1	0.373	**(m)**
x2	~~	x3	8.532	0.218	0.1	0.268	**(m)**
x3	~~	x5	7.858	-0.130	0.1	0.577	**(m)**
x1	~~	x9	7.335	0.138	0.1	0.502	**(m)**
x4	~~	x6	6.220	-0.235	0.1	0.186	**(m)**
x4	~~	x7	5.920	0.098	0.1	0.698	**(m)**
x1	~~	x7	5.420	-0.129	0.1	0.438	**(m)**
x7	~~	x9	5.183	-0.187	0.1	0.230	**(m)**
x3	~~	x9	4.126	0.102	0.1	0.515	**(m)**
x4	~~	x8	3.805	-0.069	0.1	0.806	(nm)
x1	~~	x2	3.606	-0.184	0.1	0.178	(i)
x1	~~	x4	3.554	0.078	0.1	0.673	(i)
x4	~~	x5	2.534	0.186	0.1	0.137	(i)
x2	~~	x9	1.895	0.075	0.1	0.449	(i)
x3	~~	x6	1.855	0.055	0.1	0.701	(i)
x5	~~	x7	1.233	-0.049	0.1	0.612	(i)
x5	~~	x9	0.999	0.040	0.1	0.713	(i)
x1	~~	x3	0.935	-0.139	0.1	0.107	(i)
x5	~~	x6	0.916	0.101	0.1	0.158	(i)
x2	~~	x6	0.785	0.039	0.1	0.618	(i)
x3	~~	x7	0.638	-0.044	0.1	0.435	(i)
x1	~~	x8	0.634	-0.041	0.1	0.491	(i)
x2	~~	x4	0.534	-0.034	0.1	0.581	(i)
x1	~~	x5	0.522	-0.033	0.1	0.587	(i)
x5	~~	x8	0.347	0.023	0.1	0.725	(i)
x6	~~	x8	0.275	0.018	0.1	0.839	(nm)
x6	~~	x7	0.259	-0.020	0.1	0.735	(i)
x4	~~	x9	0.196	-0.016	0.1	0.795	(nm)
x3	~~	x4	0.142	-0.016	0.1	0.662	(i)
x6	~~	x9	0.097	-0.011	0.1	0.828	(nm)
x3	~~	x8	0.059	-0.012	0.1	0.519	(i)
x2	~~	x8	0.054	-0.012	0.1	0.459	(i)
x1	~~	x6	0.048	0.009	0.1	0.713	(i)
x2	~~	x5	0.023	-0.008	0.1	0.500	(i)

Table 3: Modification Indices With Respect To Factor Loadings

Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
visual	=~	x9	36.411	0.577	0.4	0.987	*epc:m*
visual	=~	x7	18.631	-0.422	0.4	0.984	epc:nm
textual	=~	x3	9.151	-0.272	0.4	0.994	epc:nm
textual	=~	x1	8.903	0.350	0.4	0.926	epc:nm
visual	=~	x5	7.441	-0.210	0.4	0.999	epc:nm
textual	=~	x9	4.796	0.138	0.4	1.000	epc:nm
visual	=~	x8	4.295	-0.210	0.4	0.976	epc:nm
textual	=~	x8	3.359	-0.121	0.4	1.000	(nm)
visual	=~	x6	2.843	0.111	0.4	1.000	(nm)
speed	=~	x2	1.580	-0.198	0.4	0.717	(i)
visual	=~	x4	1.211	0.077	0.4	1.000	(nm)
speed	=~	x3	0.716	0.136	0.4	0.702	(i)
speed	=~	x6	0.273	0.044	0.4	0.997	(nm)
speed	=~	x5	0.201	-0.044	0.4	0.983	(nm)
textual	=~	x7	0.098	-0.021	0.4	1.000	(nm)
textual	=~	x2	0.017	-0.011	0.4	0.996	(nm)
speed	=~	x1	0.014	0.024	0.4	0.488	(i)
speed	=~	x4	0.003	-0.005	0.4	0.995	(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: 54 - 30 = 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 85.306 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 0.01 then the test is not statistically significant at 1 % 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~~ has 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 918.852 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 0.092. The upper bound of the 90% confidence interval of the RMSEA is 0.114 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 0.08 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 0.931 which is greater or equal than the threshold value 0.90, suggesting an acceptable 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 0.896 which is smaller or equal than the threshold value 0.90, suggesting a poor model fit.

↳ nearly 0.9, not so bad

Summary of the Goodness of Fit Indices

There exists support for the model from the SRMR and CFI model fit indices. We therefore assume that the model is not excellent but eventually acceptable. We proceed by diagnosing the sources of possible misspecification.

Residuals

maybe numerate the figures and refer on them ←

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 zero-order covariance relationship between the involved variables is probably underestimated:

?

Table 4: Pairs with Underestimated Covariance

	Pair(s)	
Pair 1	x1	x9
Pair 2	x3	x9
Pair 3	x7	x8

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 zero-order covariance relationship between the involved variables is probably overestimated:

?

Table 5: Pairs with Overestimated Covariance

	Pair(s)	
Pair 1	x1	x7
Pair 2	x2	x7
Pair 3	x3	x5
Pair 4	x8	x9
Pair 5	x9	x8

Depending on the sample size, the misspecification detected by the analysis of the residual covariance resp. correlation matrices can be relevant or could be in practice neglected. This is matter of subject in the next section.

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 ~~Chi~~ χ² 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". What is ~~this~~ the abbreviation for?
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.
Following modification indices with respect to factor loadings fulfill these requirements:

Table 6: Significant and Relevant MIs With Respect To Factor Loadings

	Left	Operator	Right	Modification Index	Expected Parameter Change	Delta	Power	Decision
42	visual	=~	x9	36.411	0.577	0.4	0.987	*epc:m*

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 7: Check Completely Standardized Factor Loadings

Latent Variable	Observed Variable	Factor Loading[1]	P-Value	Significant?[2]	Relevance[3]	Direction	Check
visual	x1	0.77	<0.001	Yes	**	Ok	Ok
visual	x2	0.42	<0.001	Yes	*	Ok	Ok
visual	x3	0.58	<0.001	Yes	*	Ok	Ok
textual	x4	0.85	<0.001	Yes	***	Ok	Ok
textual	x5	0.86	<0.001	Yes	***	Ok	Ok
textual	x6	0.84	<0.001	Yes	***	Ok	Ok
speed	x7	0.57	<0.001	Yes	*	Ok	Ok
speed	x8	0.72	<0.001	Yes	**	Ok	Ok
speed	x9	0.67	<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.

where?

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 8: Check Unstandardized Factor Loadings

Latent Variable	Observed Variable	Factor Loading	P-Value	Significant?[1]	Direction
visual	x1	1.00			Ok
visual	x2	0.55	<0.001	Yes	Ok
visual	x3	0.73	<0.001	Yes	Ok
textual	x4	1.00			Ok
textual	x5	1.11	<0.001	Yes	Ok
textual	x6	0.93	<0.001	Yes	Ok
speed	x7	1.00			Ok
speed	x8	1.18	<0.001	Yes	Ok
speed	x9	1.08	<0.001	Yes	Ok

[1] 5% type I error.

merge to one table (maybe over one side)

Table 9: Interpretation of Unstandardized Factor Loadings

Interpretation of Unstandardized Factor Loadings
x1 is marker variable for visual
A 1-unit increase in visual leads to a 0.55 -unit increase in the x2
A 1-unit increase in visual leads to a 0.73 -unit increase in the x3
x4 is marker variable for textual
A 1-unit increase in textual leads to a 1.11 -unit increase in the x5
A 1-unit increase in textual leads to a 0.93 -unit increase in the x6
x7 is marker variable for speed
A 1-unit increase in speed leads to a 1.18 -unit increase in the x8
A 1-unit increase in speed leads to a 1.08 -unit increase in the x9

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 0.47 which we regard as a proof of a reasonable discriminant validity.

Factor Reliability

Error Variances

Table 10: Completely Standardized Error Variances and Communality

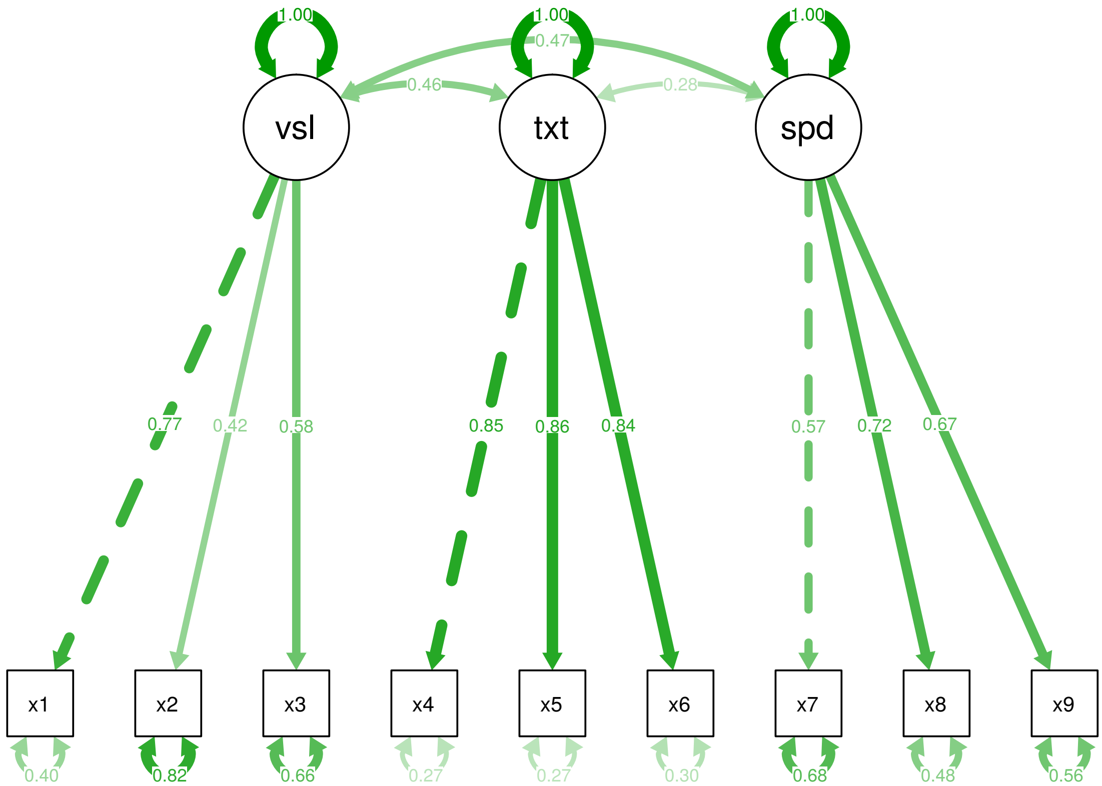
Observed Variable	Error Variance[1]	Communality[23]	P-Value	Significant Error Variance?[4]
x1	0.40	0.60	<0.001	Yes
x2	0.82	0.18	<0.001	Yes
x3	0.66	0.34	<0.001	Yes
x4	0.27	0.73	<0.001	Yes
x5	0.27	0.73	<0.001	Yes
x6	0.30	0.70	<0.001	Yes
x7	0.68	0.32	<0.001	Yes
x8	0.48	0.52	<0.001	Yes
x9	0.56	0.44	<0.001	Yes

[1] Can be interpreted as proportion of unexplained variance by the latent factor(s) (%).
[2] Corresponds to the squared factor loading.
proportion of what in what?
[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.

Path Diagram



References