Confirmatory factor analysis (CFA)

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

General remarks

1. Characteristics of CFA

- Theory-guided procedure

  - possible outcomes are limited and known in beforehand
  - new knowledge in the sence of an *accidential finding* is not possible
  - the procedure is suitable for investigating pre-specified hypotheses / research questions

General remarks

1. Characteristics of CFA

- Theory-guided procedure
- Applicable for many purposes
  - scale construction
  - scale validation
  - model comparison
  - multi-sample investigations
  - etc.

General remarks

1. Characteristics of CFA

- Theory-guided procedure
- Applicable for many purposes
- Relates manifest to latent variables

  - uses manifest information as input
  - provides parameter estimates obtained at latent level

General remarks

1. Characteristics of CFA

- Theory-guided procedure
- Applicable for many purposes
- Relates manifest to latent variables
- Provides complete accounts of data
  - all parts of the model must be correct
  - matrix is reproduced for a check

General remarks

1. Characteristics of CFA

- Theory-guided procedure
- Applicable for many purposes
- Relates manifest to latent variables
- Provides complete account of data
- Enables model evaluation
- Enables model comparison

General remarks

2. Reasons for the switch from EFA to CFA were …

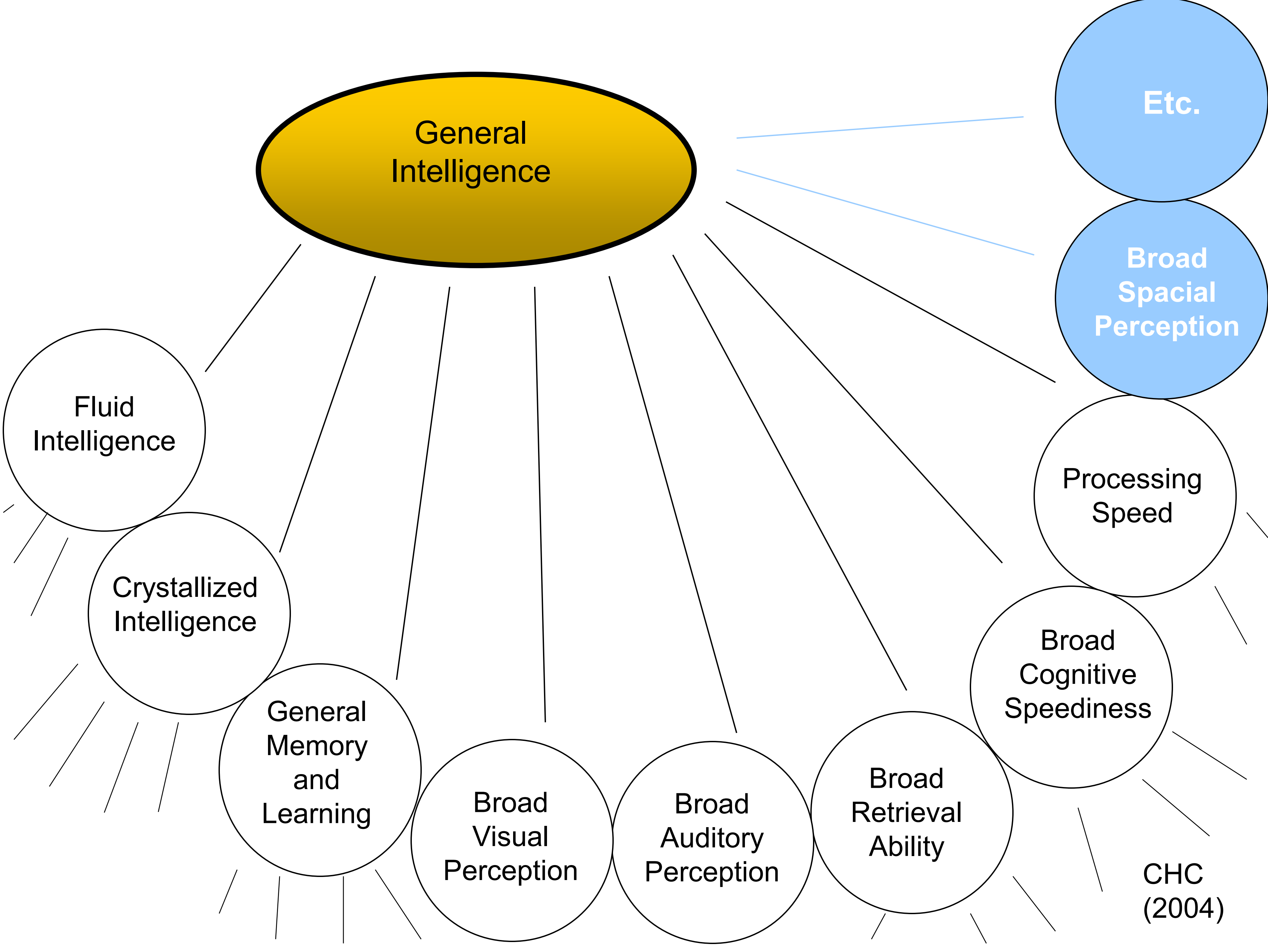
- avoidance of weaknesses of EFA:
  - > dependency of results on *selection of items*
  - > dependency of results on *sample*
  - > dependency of results on *extraction method*
  - > dependency of results on *rotation method*
- *problems in integrating research results leading to …*
  - > several differing models of personality
  - > several differing models of ability

General remarks

2. Reasons for the switch from EFA to CFA were ...

Examples from intelligence research:

- *Spearman (1904) proposes a model of general intelligence*
- *Thurstone (1938) proposes non-hierarchical multi-factor model*
- *Cattell (1941) proposes a two-factor model*
- *Guilford (1967) proposes a model with 120 factors*
- *Carroll (1993) proposes a hierarchical model with three hierarchical levels*
- *...*
- *...*


General Intelligence
Etc.
Broad Spacial Perception
Fluid Intelligence
Processing Speed
Crystallized Intelligence
Broad Cognitive Speediness
General Memory and Learning
Broad Visual Perception
Broad Auditory Perception
Broad Retrieval Ability
CHC (2004)

General remarks

3. Most important improvements (from EFA to CFA) for applied research:

- ….. the possibility to apply the *same model* (simultaneously) to *several samples*
- ….. the possibility to compare *several models* with respect to the *same sample*

General remarks

4. Confirmatory factor analysis means employing *several* methods in investigating data *successively* .

Several of them are considered in this course unit: $\longrightarrow$ the steps

General remarks: *the steps*

1. Formation of hypotheses
2. Construction of path diagram
3. Formalization of model structure
4. Check of identification
5. Selection of a research strategy
6. Parameter estimation
7. Evaluation of results

Outline

Confirmatory factor analysis

- General remarks
- *Formal similarity of EFA and CFA*
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

Formal similarity of EFA and CFA

- This section starts from the fundamental theorem of factor analysis to make the relationship of EFA and CFA obvious

$$z_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p}\left(\lambda_{ik}F_{mk}\right) + \varepsilon_{mi}$$

- The theorem is gradually transformed into a model of measurement

Formal similarity of EFA and CFA

- ...

$$z_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p}\left(\lambda_{ik}F_{mk}\right) + \varepsilon_{mi}$$

- 1. Standardization is replaced by centering

$$y_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi}$$

Formal similarity of EFA and CFA

- …

$$z_{mi} = \lambda_{i1} F_{m1} + \lambda_{i2} F_{m2} + \ldots + \lambda_{ip} F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p} \left(\lambda_{ik} F_{mk} \right) + \varepsilon_{mi}$$

- 2. The number of factors is reduced to one (customary version)

$$y_{mi} = \lambda_{i1} F_{m1} + \varepsilon_{mi}$$

… more factors are possible

Formal similarity of EFA and CFA

- …

$$z_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p}\left(\lambda_{ik}F_{mk}\right) + \varepsilon_{mi}$$

- 3. The equations for all items are considered simultaneously

$$y_{m1} = \lambda_{11}F_{m1} + \varepsilon_{m1}$$

$$\vdots$$

$$y_{mi} = \lambda_{i1}F_{m1} + \varepsilon_{mi}$$

$$\vdots$$

$$y_{mp} = \lambda_{p1}F_{m1} + \varepsilon_{mp}$$

Formal similarity of EFA and CFA

- …

$$z_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p}\left(\lambda_{ik}F_{mk}\right) + \varepsilon_{mi}$$

- 3. The equations are combined to give vectors

$$\mathbf{y}_{\mathrm{m}} = \boldsymbol{\lambda}F_{m} + \boldsymbol{\varepsilon}_{m}$$

Formal similarity of EFA and CFA

- …

$$z_{mi} = \lambda_{i1}F_{m1} + \lambda_{i2}F_{m2} + \ldots + \lambda_{ip}F_{mp} + \varepsilon_{mi} = \sum_{k=1}^{p}\left(\lambda_{ik}F_{mk}\right) + \varepsilon_{mi}$$

- 4. The factor score (F) is replaced by the latent variable to obtain a general version (= a version that is independent of a specific person)

$$\mathbf{y} = \boldsymbol{\lambda}\ \xi + \boldsymbol{\varepsilon}$$

… both F and ξ characterize a person property

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- *Formation of hypotheses*
- Construction of path diagram
- Check of identification
- An alternative model
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

Formation of hypotheses

Considering constructs of different degrees of complexity:

- **it may be assumed to consist in *one unit only***

  e.g. altruism

- **it may show a specific structure including *several units***

  e.g. optimism may include separate optimism and pessimism sub-units

Formation of hypotheses

Considering constructs of different degrees of complexity:

…

Possible structures of optimism:

1. Optimism as one dimension (that is independent of the pessimism dimension)
2. … is one dimension with two poles: optimism and pessimism
3. Optimism as a hierarchical structure: one independent upper-level dimension and two lower-level dimensions of optimism and pessimism
4. Optimism as main dimension associated with a method factor (e.g. social desirability)

Formation of hypotheses

Considering constructs of different degrees of complexity:

- In the case of a complex construct the representation by a *single* hypothesis is insufficient.

- In such a case <u>several related hypotheses</u> have to be worked out and investigated.

Formation of hypotheses

Formation regarding a complex construct: e.g. *affectivity*

(Structure-of-construct) Hypothesis 1:

- e.g. affectivity is composed of two sub-units (positive and negative ones) that are related to each other

Formation of hypotheses

Formation regarding a complex construct: e.g. *affectivity*

(Structure-of-construct) Hypothesis 1:

- e.g. affectivity is composed of two sub-units (positive and negative ones) that are related to each other

(Measurement) Hypothesis (sub-unit 1) 2:

- positive affectivity is represented by three manifest variables < to be specified >

(Measurement) Hypothesis (sub-unit 2) 3:

- negative affectivity is represented by three manifest variables < to be specified >

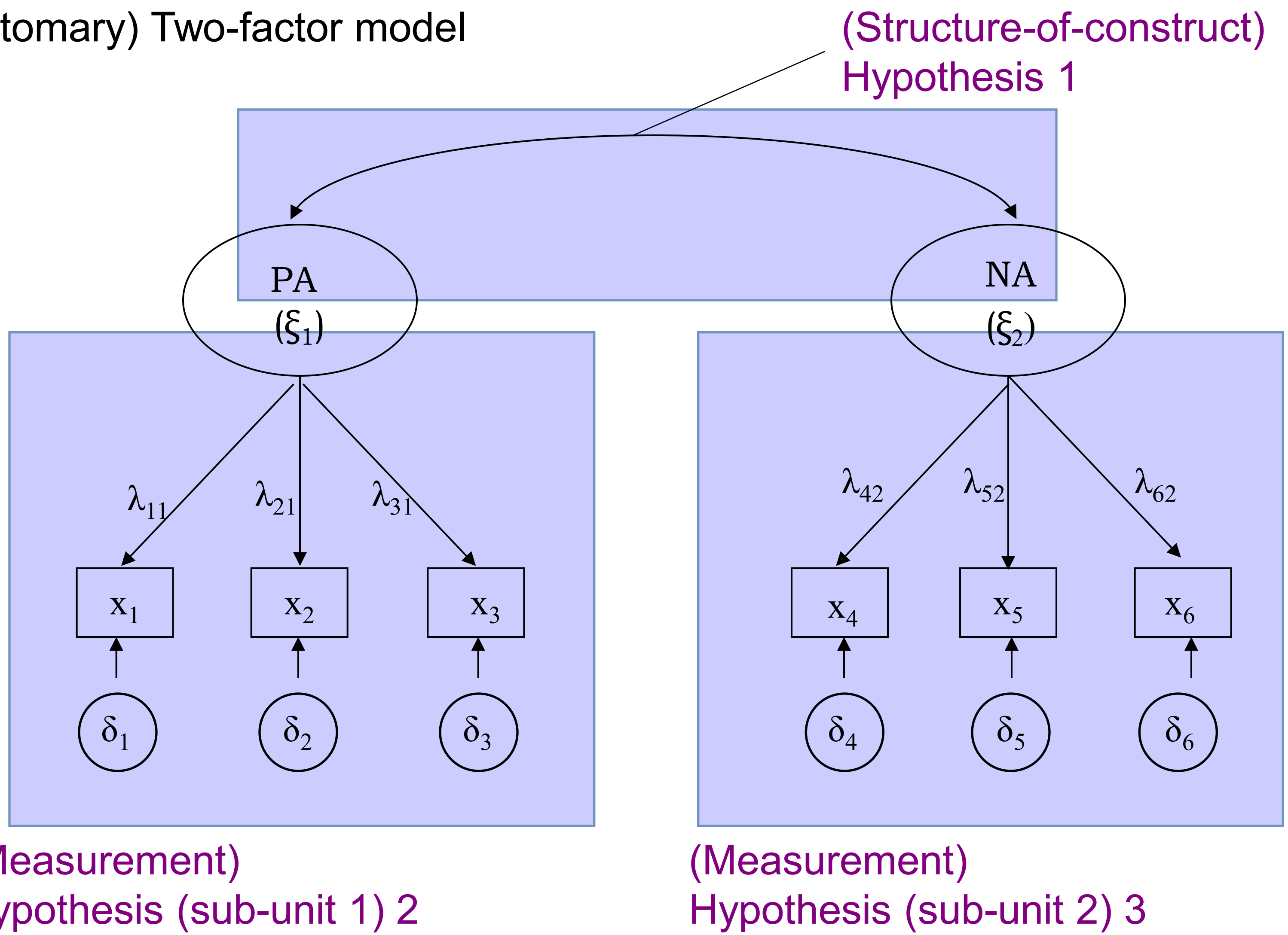
Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- *Construction of path diagram*
- Check of identification
- An alternative model
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

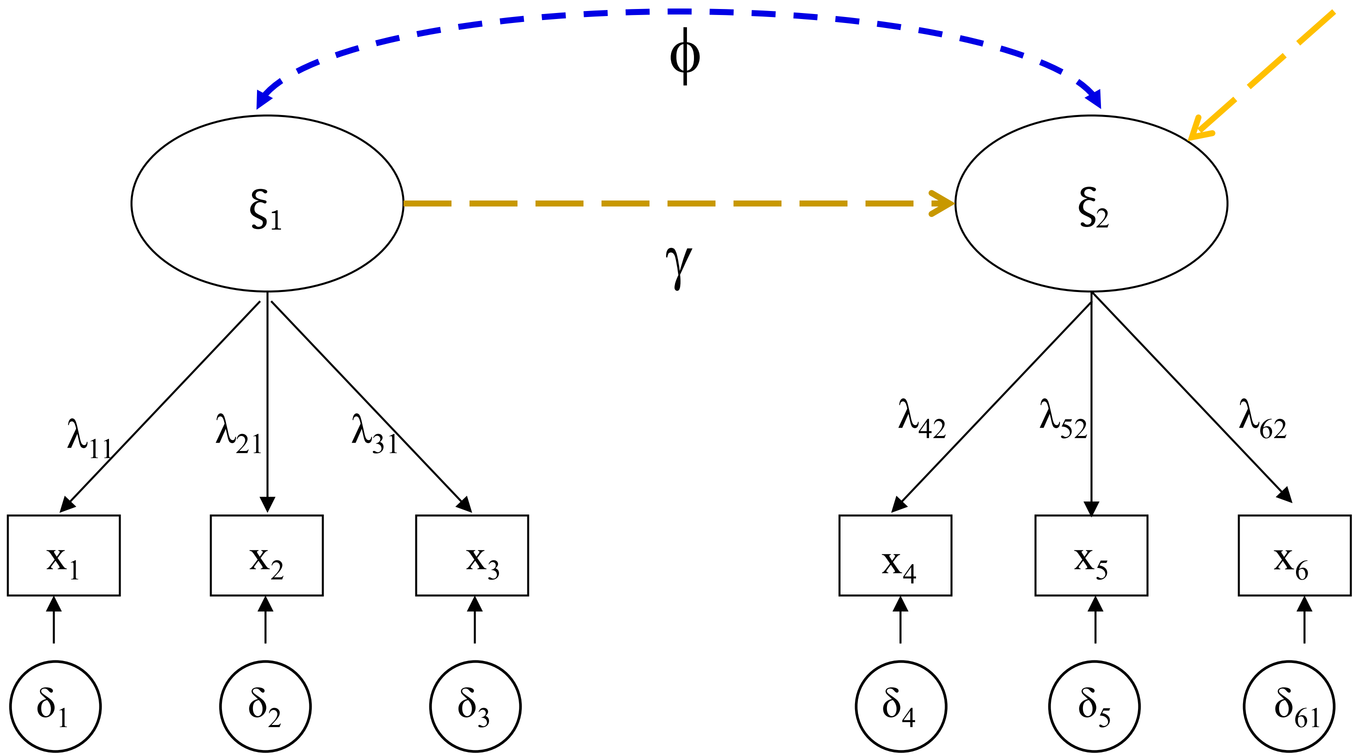
Construction of path diagram

(customary) Two-factor model



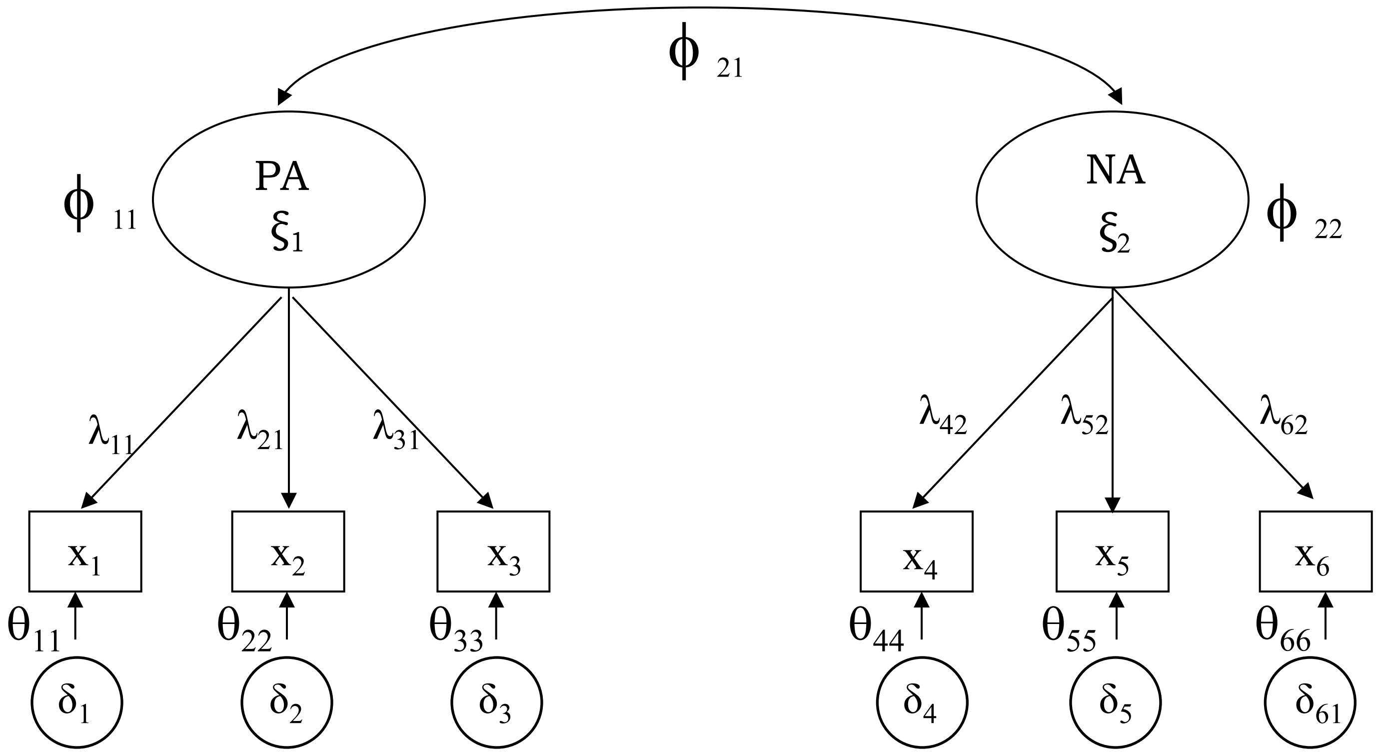
Construction of path diagram

Similarity of complex measurement model and full structural equation model:



Construction of path diagram

Preparation for data analysis:



Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- *Check of identification*
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

Check of identification

Possible outcomes of the identification check:

- The model is **not identified** if there are more parameters than items of information

- The model is **just identified (identified)** if the number of parameters corresponds to the number of items of information

- The model is **identified (over-identified)** if the number of items of information is larger than the number of parameters

Check of identification

df = available information (*s*) – number of parameters (*t*)

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	s_{11}					
x_2	s_{21}	s_{22}				
x_3	s_{31}	s_{32}	s_{33}			
x_4	s_{41}	s_{42}	s_{43}	s_{44}		
x_5	s_{51}	s_{52}	s_{53}	s_{54}	s_{55}	
x_6	s_{61}	s_{62}	s_{63}	s_{64}	s_{65}	s_{66}

The items of information for computing the ***df***!

Check of identification

Items of information included in the variance-covariance matrix

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	s_{11}					
x_2	s_{21}	s_{22}				
x_3	s_{31}	s_{32}	s_{33}			
x_4	s_{41}	s_{42}	s_{43}	s_{44}		
x_5	s_{51}	s_{52}	s_{53}	s_{54}	s_{55}	
x_6	s_{61}	s_{62}	s_{63}	s_{64}	s_{65}	s_{66}

$$s = \frac{p(p+1)}{2} \qquad = 21$$

p: number of manifest variables

Check of identification

- The model is **(over-) identified** if the number of items of information is larger that the number of parameters

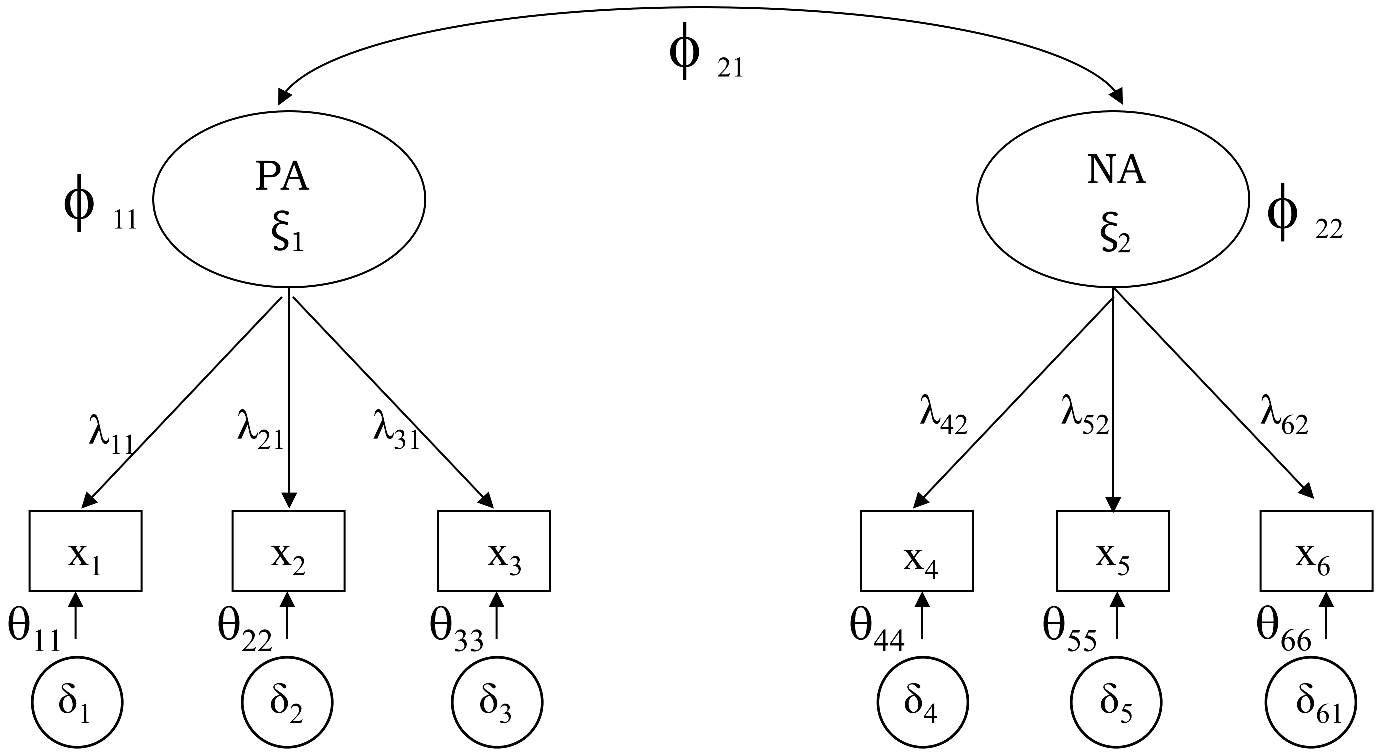
This is the desirable situation for modeling latent variables

Practice

Please, check whether the (whole) affectivity model is identified !

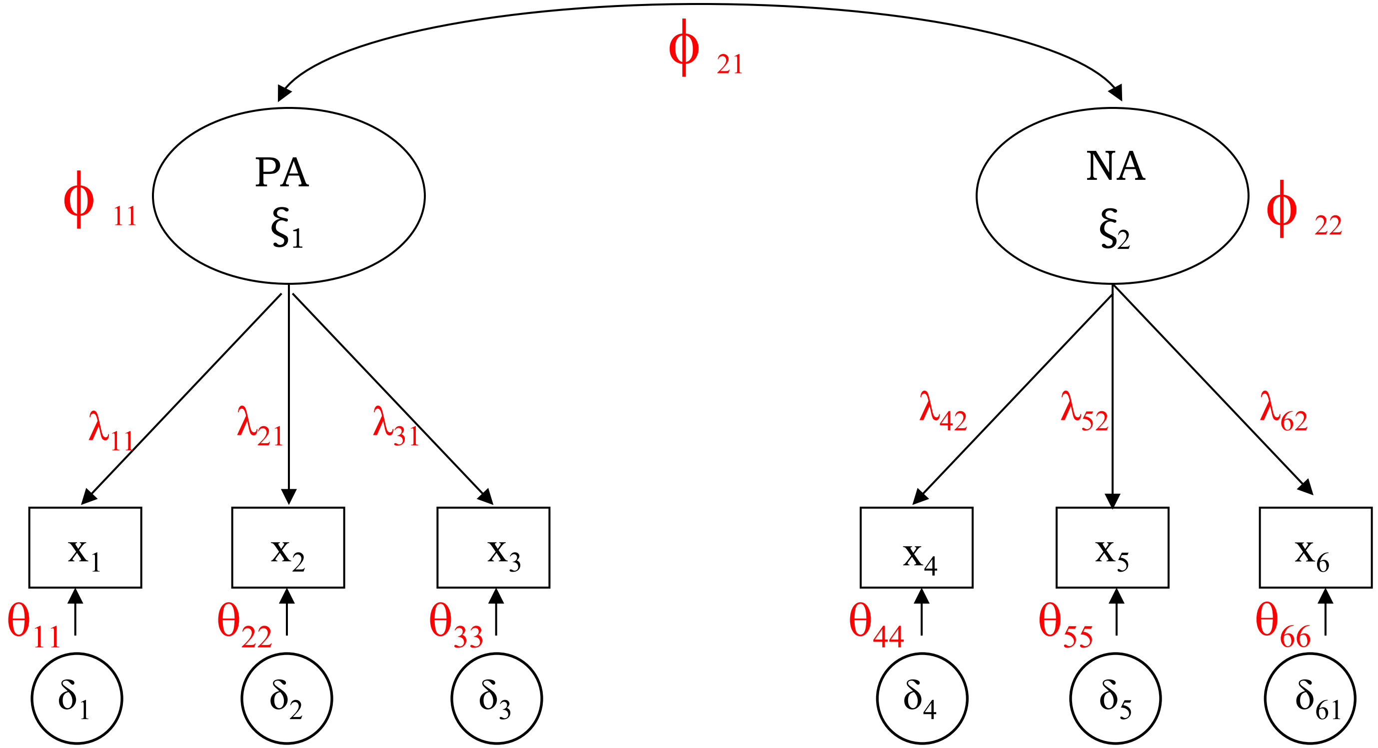
What is the degree of freedom?

Model



ϕ_{21}
PA
ξ_1
ϕ_{11}
NA
ξ_2
ϕ_{22}
λ_{11}
λ_{21}
λ_{31}
λ_{42}
λ_{52}
λ_{62}
x_1
x_2
x_3
x_4
x_5
x_6
θ_{11}
θ_{22}
θ_{33}
θ_{44}
θ_{55}
θ_{66}
δ_1
δ_2
δ_3
δ_4
δ_5
δ_{61}

Model



Number of parameters is *15*

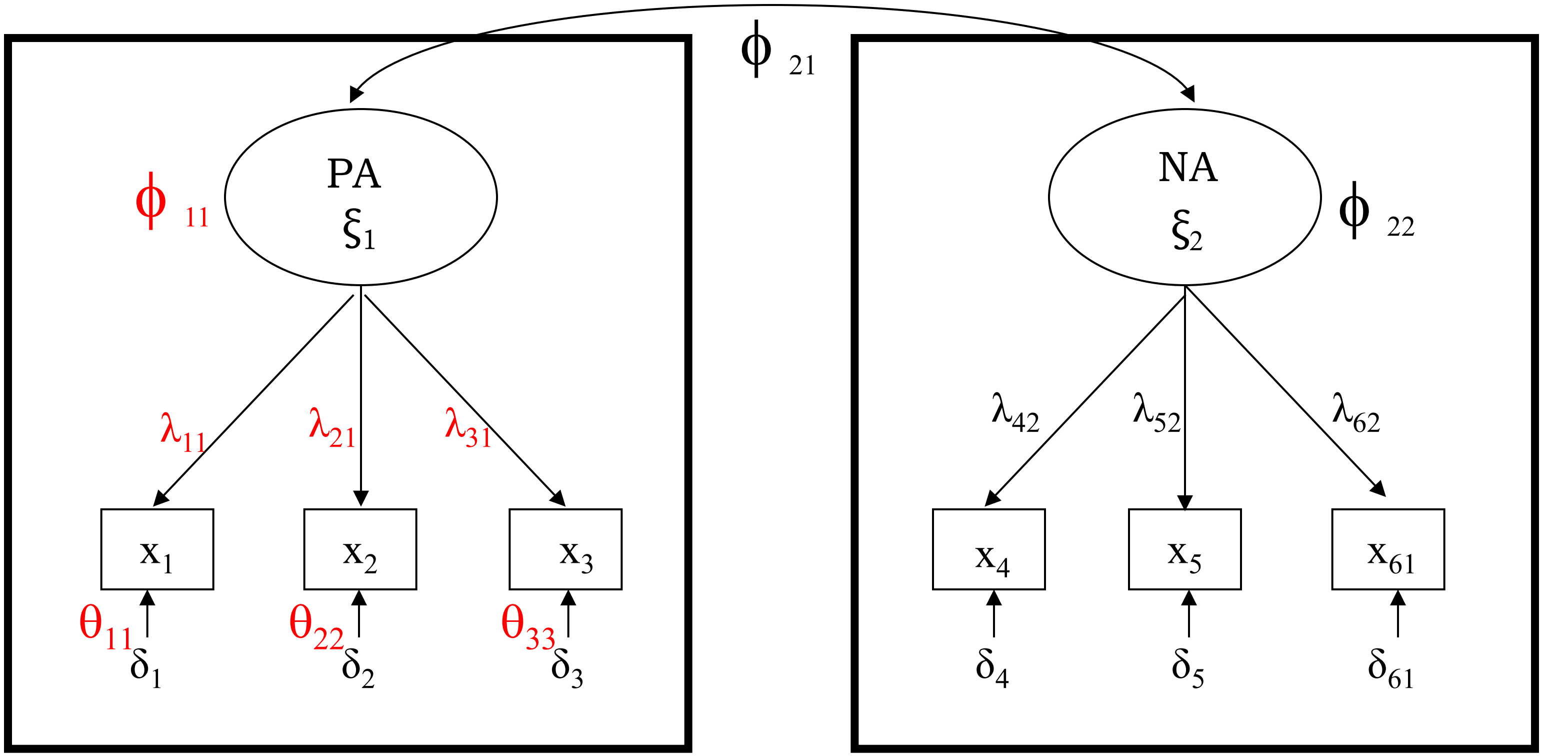
Check of identification

- Items of information
  - covariance matrix: 21

- Number of parameters: 15

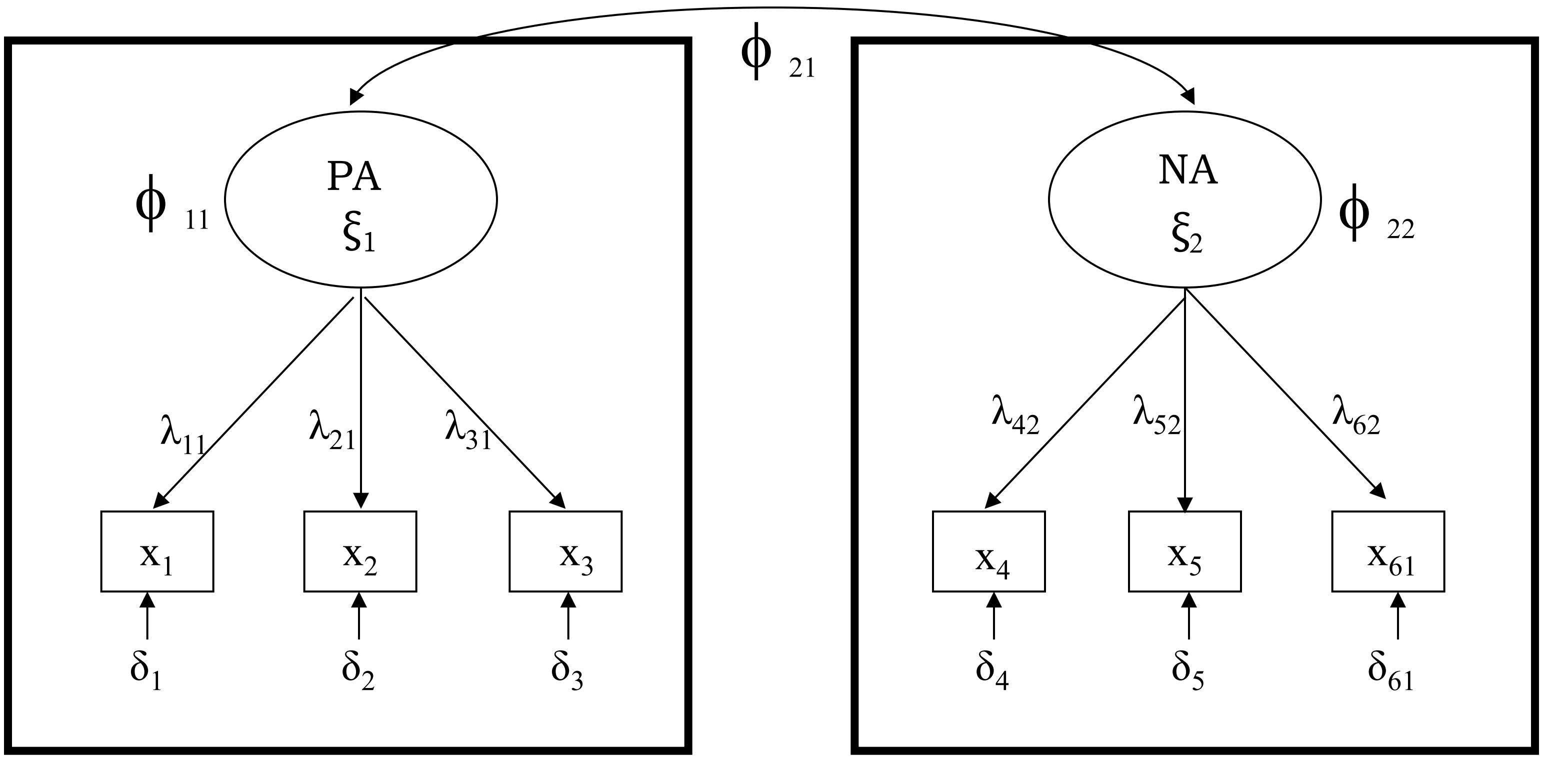
- df covariance = 21- 15 = 6
  …minus 2 for scaling = 4

Problem: models of measurement



Information number is 6 (covariances)
Number of parameters is 7

Problem: models of measurement



What is to be done to assure that the models are (***over***) ***identified***?

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- *On the research strategy*
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

The research strategy

The research strategy

Regarding the research strategy, there has been a change in the last 20 years:

There was a switch from the old to the new (modern) strategy.

Old research strategy

The *old* research strategy was focused on **one** model only!

If this model did not fit, it was modified until there was good model fit.

Old research strategy

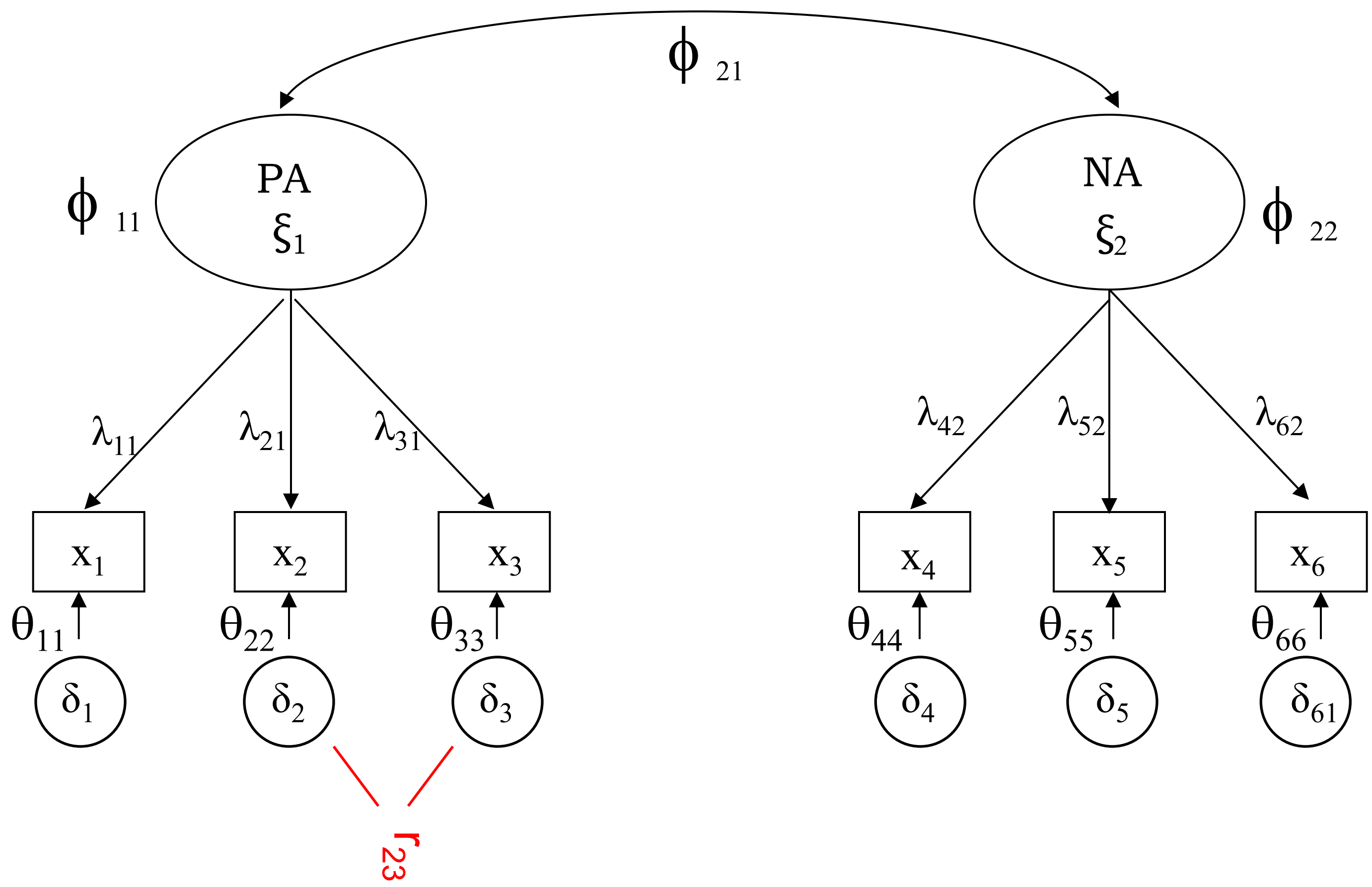
The *old* research strategy was focused on <u>**one**</u> model only. This means ...

- the fit of this model was checked.
- if the fit proved to be good, no further action was necessary.
- if the fit was <u>not good</u>, *modification indices* were consulted.
- the most promising option was realized.
- the check was repeated, until ... the fit was good.

Old research strategy

An example of model modification

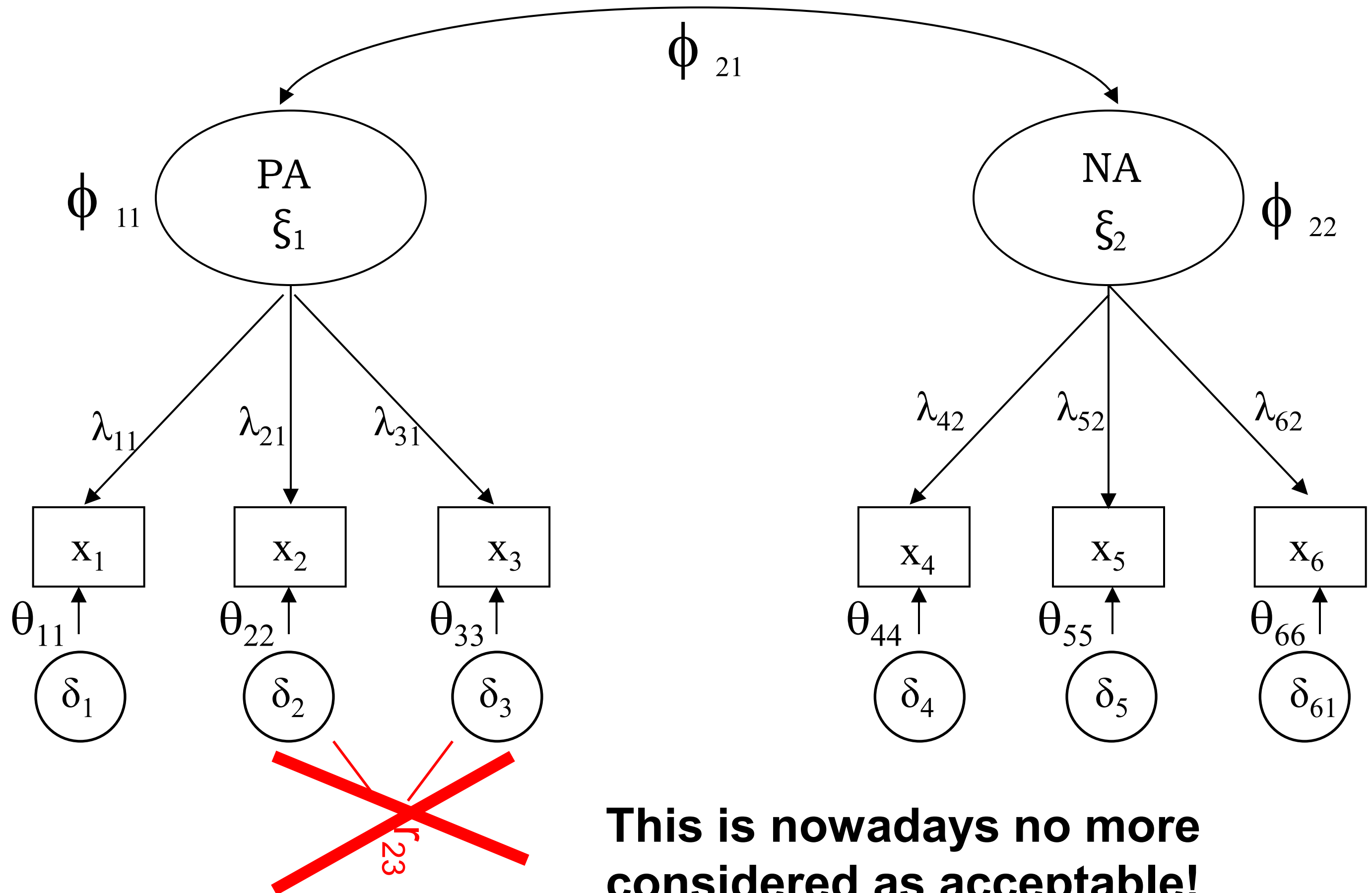
Assume that the program proposes a ***correlation*** *between* δ_2 *and* δ_3



Old research strategy

An example of model modification

*Assume that the program proposes a **correlation** between δ_2 and δ_3*

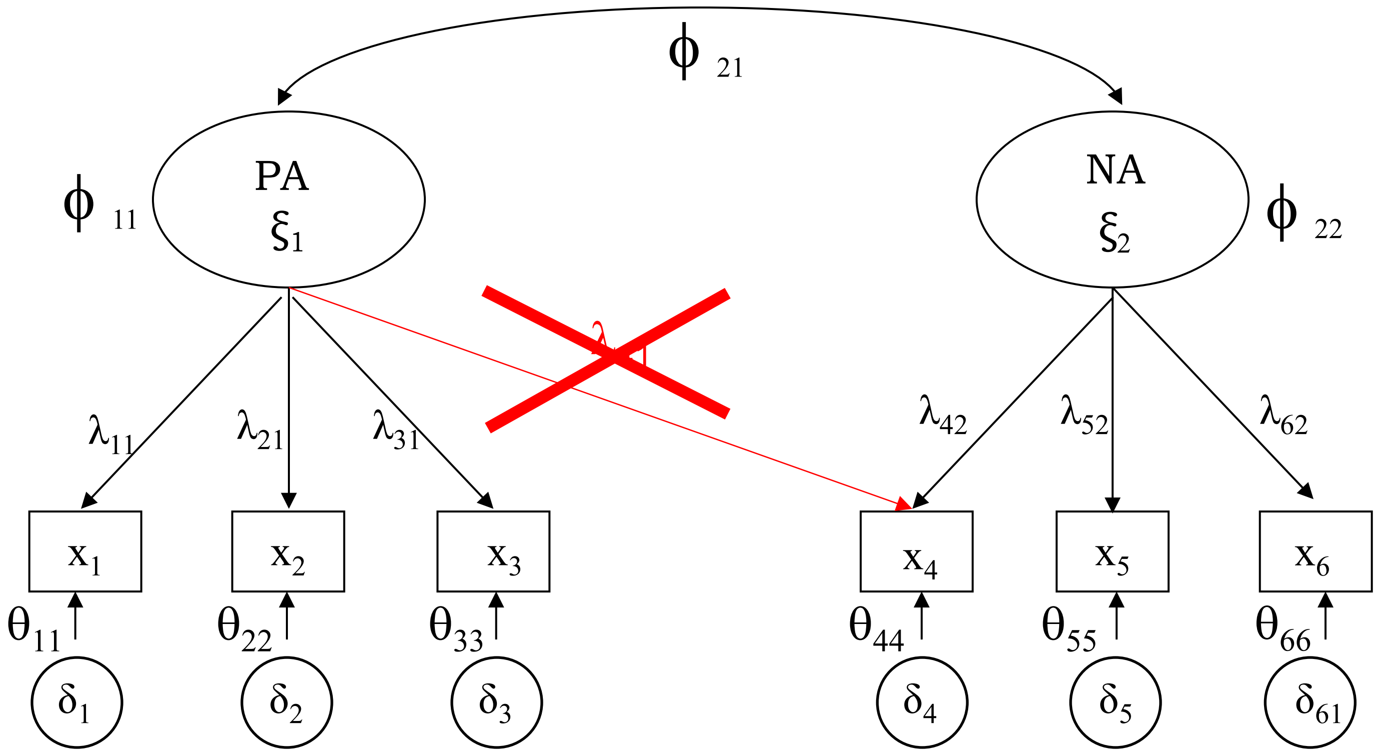


This is nowadays no more considered as acceptable!

Old research strategy

An example of model modification

*Assume that the program proposes a **coss-loading** of x_4*



Old research strategy

The *old* research strategy was focused on **<u>one</u>** model only!

If this model did not fit, it was modified until there was good model fit.

Now considered as …….. ad-hoc modifications without theoretical justification

The *new* research strategy

The research strategy is based on the following basic considerations:

- each model is *wrong*
- „best“ does not automatially mean „true“ oder „correct“
- being the „best“ model with respect to one data set does not guarantee that this model is also best with respect to other data sets
- nevertheless, the <u>*best model*</u> out of the set of wrong models has to be identified

The research strategy

Starting from these considerations nowadays there is the following general strategy:

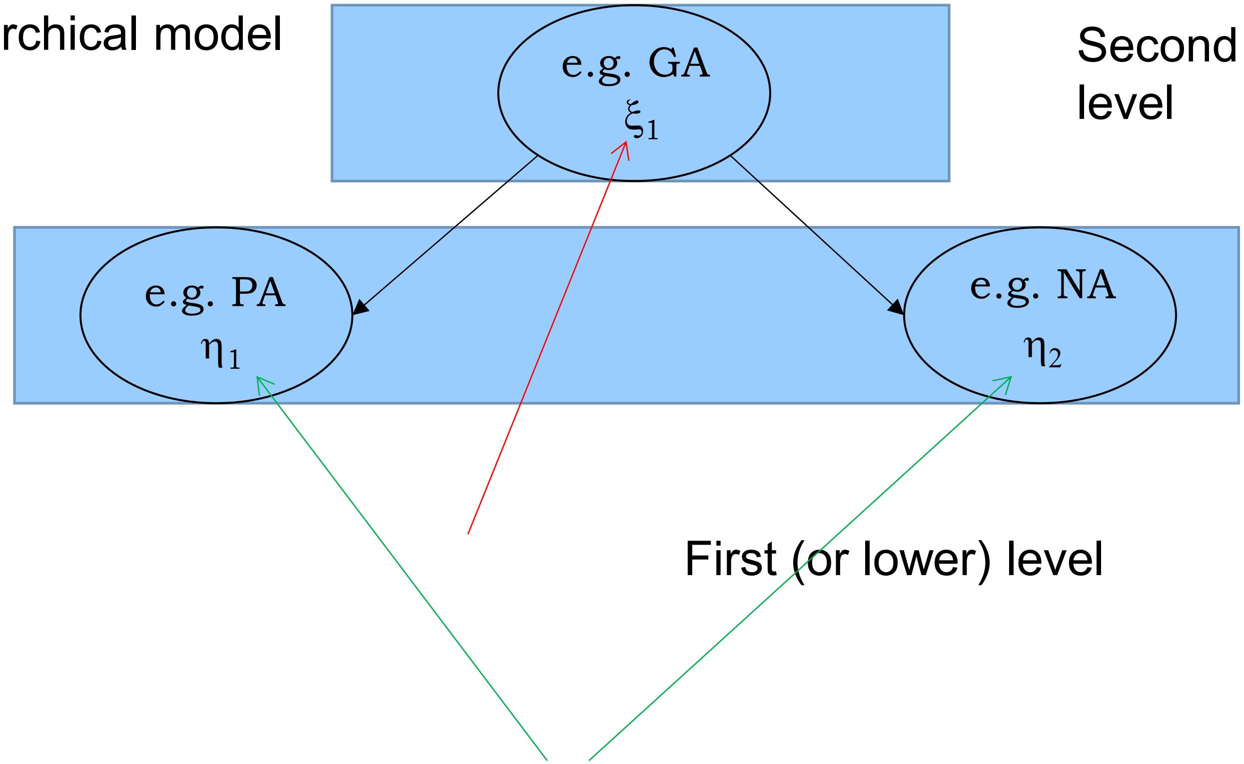
- In the first step **several** models that appear to be reasonable (out of various reasons) are specified

For the affectivity scale this strategy is likely to lead to the following set of models:

-a one-factor model with a PA-NA latent variable
-a two-factor model
-a hierarchical model with two levels (PA-NA on the top level)
-a bifactor model

The research strategy

The hierarchical model

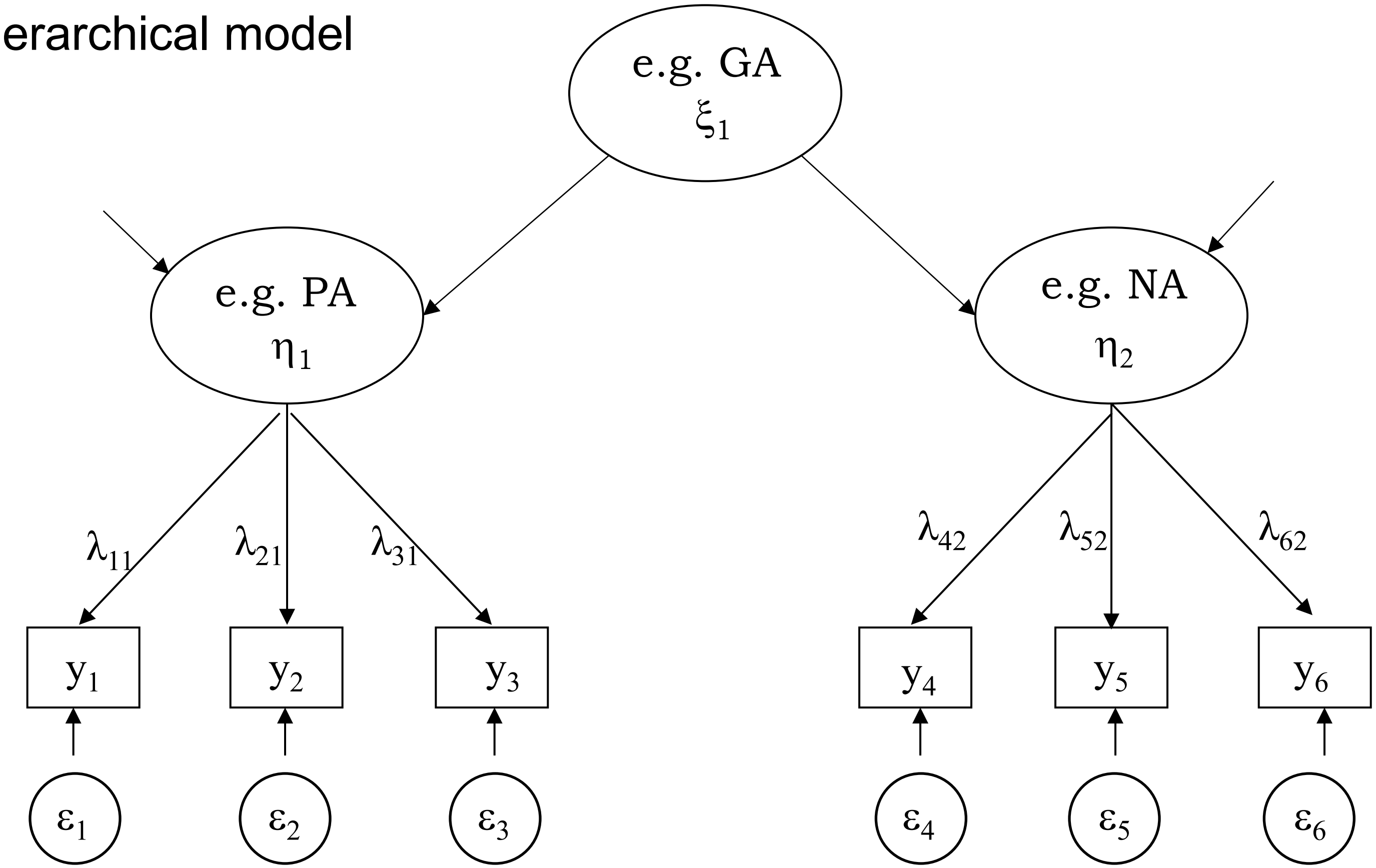


Second (or upper) level

First (or lower) level

The research strategy

The hierarchical model



The research strategy

The bifactor model

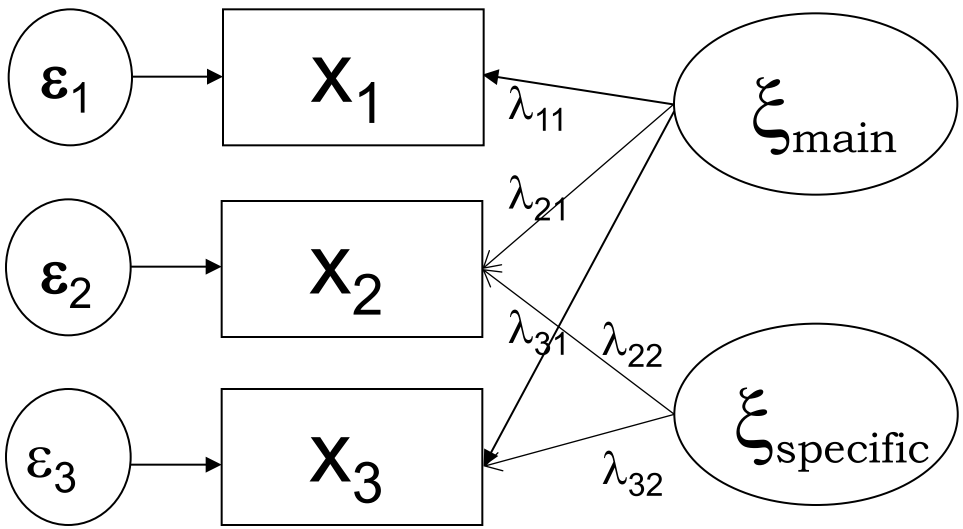
... is a model with two factors that has recently become populare as replacement of the customary two-factor model.

The research strategy

- The bifactor model is a model with two factors.
- One factor is a general factor and the other one a specific factor.
- The specific factor is *nested* within the general factor.

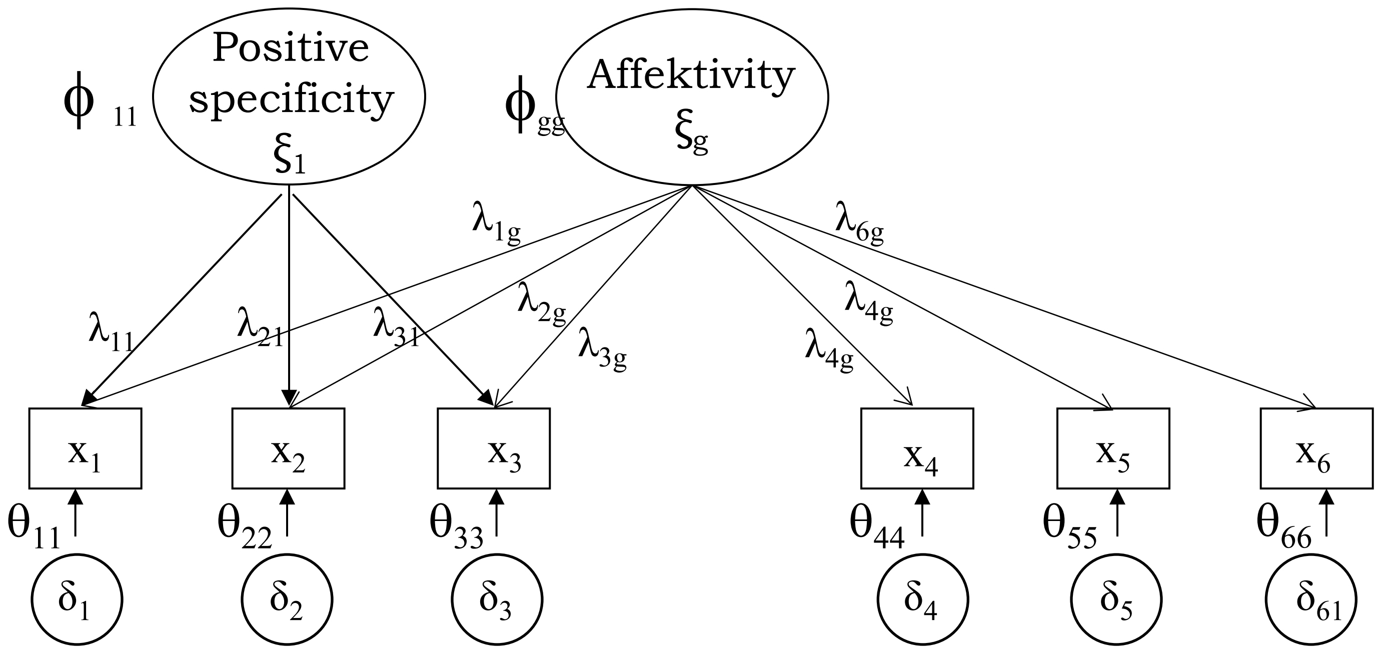
The research strategy

Path diagram of a bifactor model



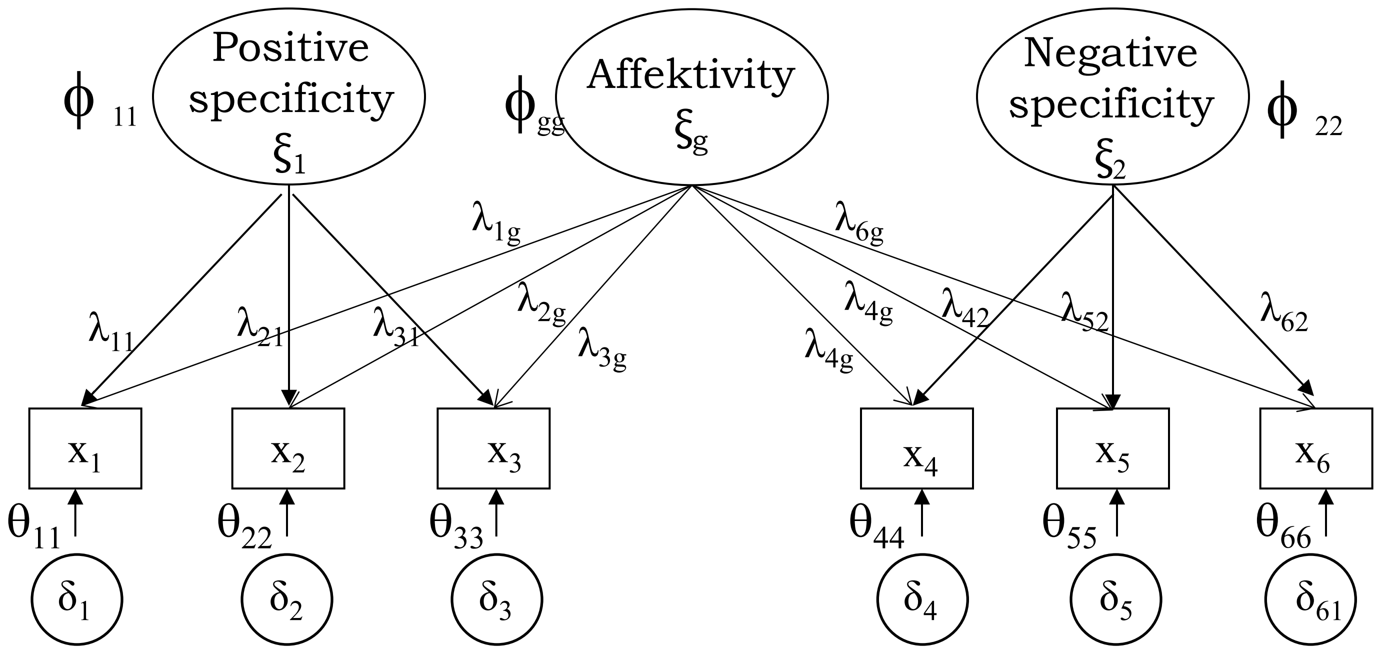
The research strategy

The affectivity model as bifactor model:



The research strategy

The affectivity model as extended bifactor model:



The research strategy

Starting from these considerations nowadays there is the following strategy:

- In the first step <u>several</u> models that appear to be reasonable (out of various reasons) are specified
- After computing the fit statistics, the fit results of the models are compared with each other
- The model that shows the <u>best fit</u> (and is simpler than other models showing similar model fit) is considered as the „best“ model.

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- *Parameter estimation*
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

Preparation for estimation

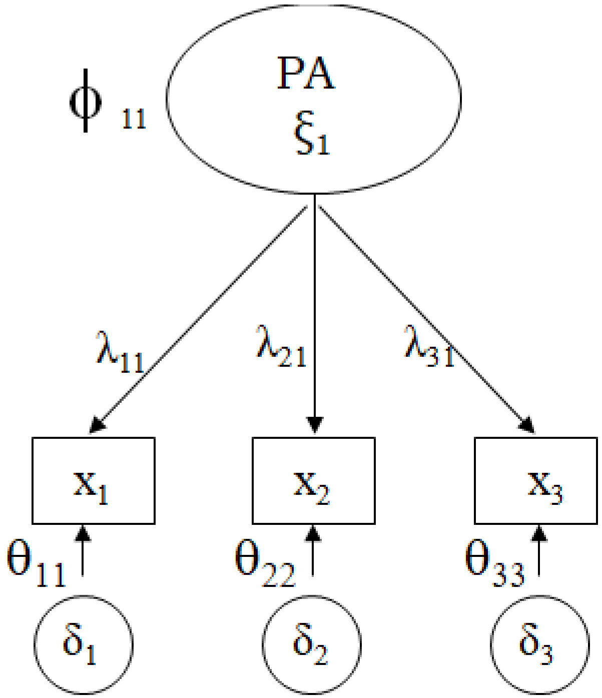
The information on the to-be-estimated parameters needs to be prepared for the computer run.

i.e. for all parameters it must be stated whether they are

- free for estimation
- fixed to zero
- set equal to a constant

Preparation for estimation

An example:



LISREL
Notation

VA 1 => ϕ_{11}	VA 1 PH 1 1
Free => λ 1 1	FR LX 1 1
Free => λ 2 1	FR LX 2 1
…………	……….
Fixed 0 => θ 1 1	FI TD 1 1
Free => θ 2 2	FR TD 2 2
…………	

Parameter estimation

- … is mostly conducted by means of the maximum likelihood method

- … is conducted for finding the parameter values for the model that enable a very good agreement of the covariance model with the empirical variances and covariances

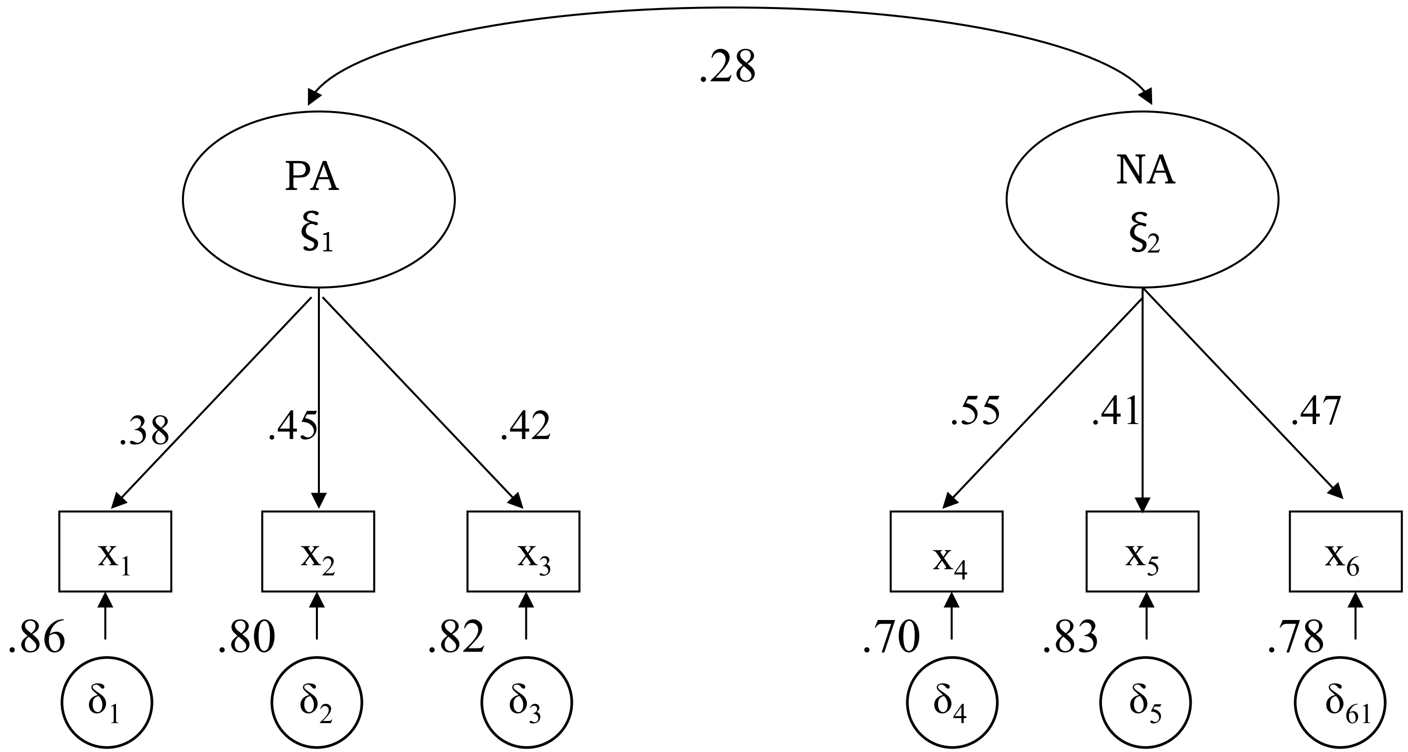
Parameter estimation

- … is mostly conducted by means of the maximum likelihood method

- Parameter estimation mostly uses an iterative algorithm: (mostly) the *expectation-maximization algorithm*

Parameter estimates

The outcome (standardized estimates) – an example:



Parameter estimates

Supplement

There is the possibility …

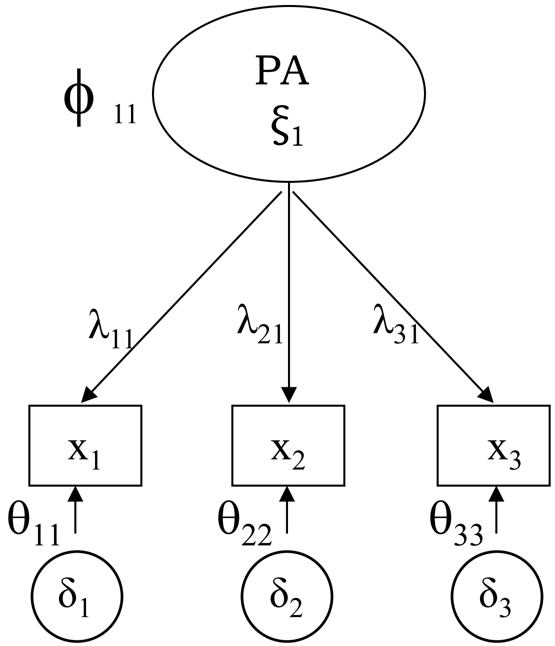
… to use of the diagram with parameters for checking *whether the parameter estimates are correct*

… the variance of the manifest variable: **var(x) =** $\lambda^2\phi+\theta$

Parameter estimates

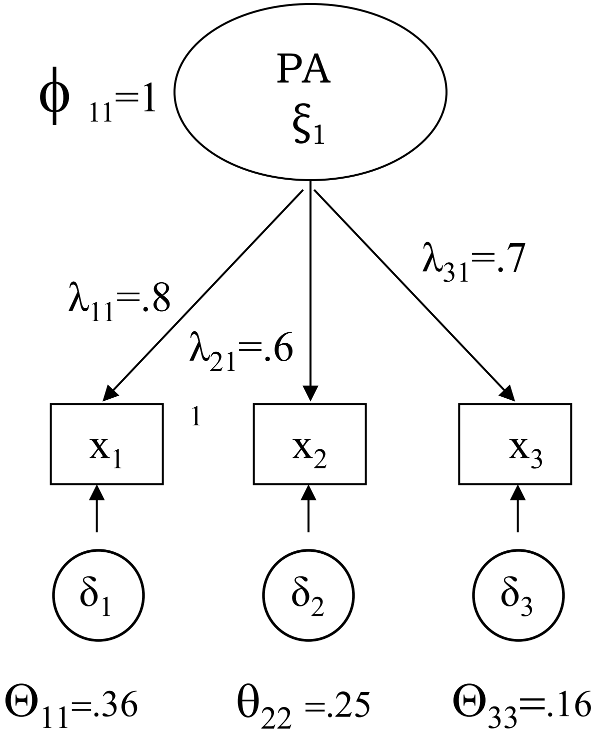
Use of the diagram with parameters for checking *whether the parameter estimates are correct* :

… the variance of the manifest variable: $\mathbf{var(x_1) = \lambda_{11}{}^2\phi_{11}+\theta_{11}}$



Parameter estimates

Use of the diagram with parameters for checking *whether the parameter estimates are correct* :



… the variance of the manifest variable: $\mathbf{var(x_1) = \lambda_{11}^2\phi_{11}+\theta_{11}}$

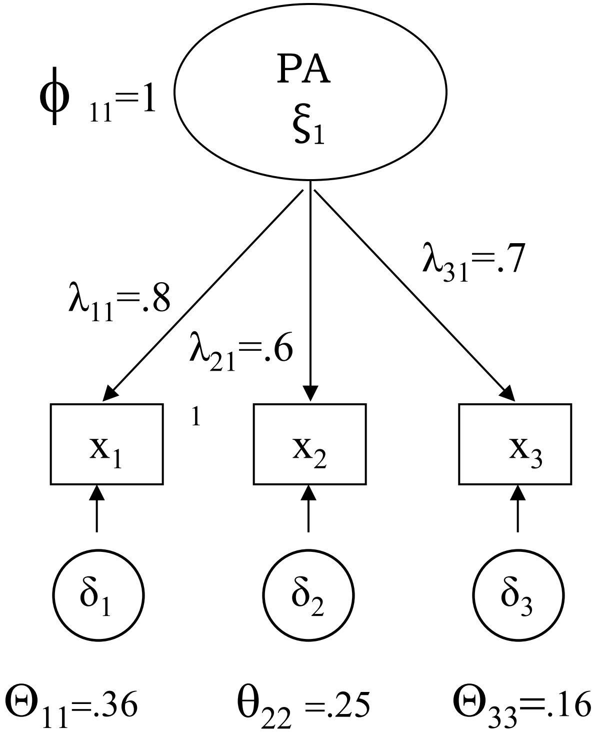
$var(x_1) = .8^2 \times 1.0 + .36$

Result: 1.0

Check:
$$\begin{vmatrix} 1.00 & XX & XX \\ XX & ?? & XX \\ XX & XX & ?? \end{vmatrix}$$

Parameter estimates

Use of the diagram with parameters for checking *whether the parameter estimates are correct* :



… the variance of the manifest variable: $\mathbf{var(x_1) = \lambda_{11}^2\phi_{11}+\theta_{11}}$

Practice 1: compute var($\mathbf{x_2}$)

Result: .61

Practice 2: compute var($\mathbf{x_3}$)

Result: .65

Parameter estimates

Supplement

There is the possibility …

… to use of the diagram with parameters for checking *whether the parameter estimates are correct*

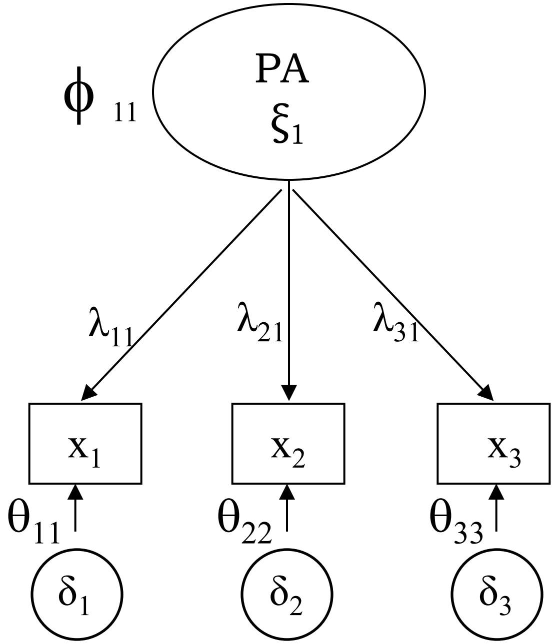
… the variance of the manifest variable: **var(x) =** $\lambda^2\phi+\theta$

… the **covariance** of manifest variables: **cov(x₁,x₂) =** $\lambda_{11}\phi_{11}\lambda_{21}$

Parameter estimates

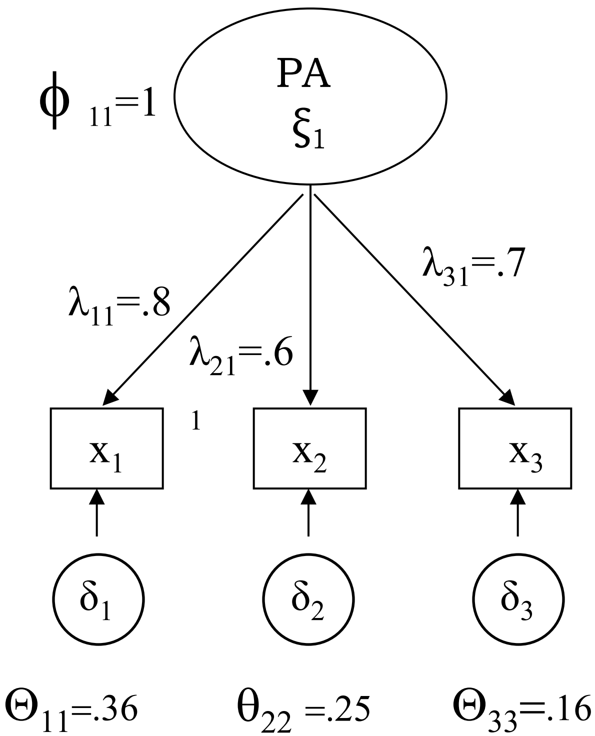
Use of the diagram with parameters for checking *whether the parameter estimates are correct* :

… the **covariance** of manifest variables: **cov(x_1,x_2) = $\lambda_{11}\phi_{11}\lambda_{21}$**



Parameter estimates

Use of the diagram with parameters for checking *whether the parameter estimates are correct* :



… the **covariance** of manifest variables: $\mathbf{cov(x_1,x_2) = \lambda_{11}\phi_{11}\lambda_{21}}$

$cov(\mathbf{x_1},\mathbf{x_2}) = .8 \times 1.0 \times .6$

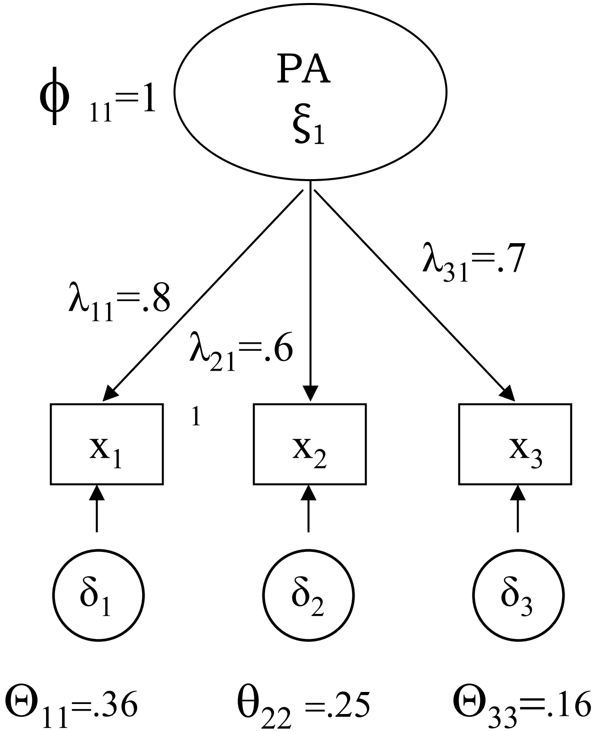
Result: .48

Check:

$$\begin{vmatrix} 1.00 & XX & XX \\ .48 & ?? & XX \\ XX & XX & ?? \end{vmatrix}$$

Parameter estimates

Use of the diagram with parameters for checking *whether the parameter estimates are correct* :



… the **covariance** of manifest variables: $\mathbf{cov(x_1,x_2) = \lambda_{11}\phi_{11}\lambda_{21}}$

Practice 3: compute cov($\mathbf{x_1}$,$\mathbf{x_3}$)

Result: .56

Practice 4: compute cov($\mathbf{x_2}$,$\mathbf{x_3}$)

Result: .42

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- *Scaling the variance parameter*
- Evaluation of model fit
- Model types
- Comparison of EFA and CFA

Scaling the variance parameter

Obtaining an estimate of the **variance** of the latent variable requires **scaling**

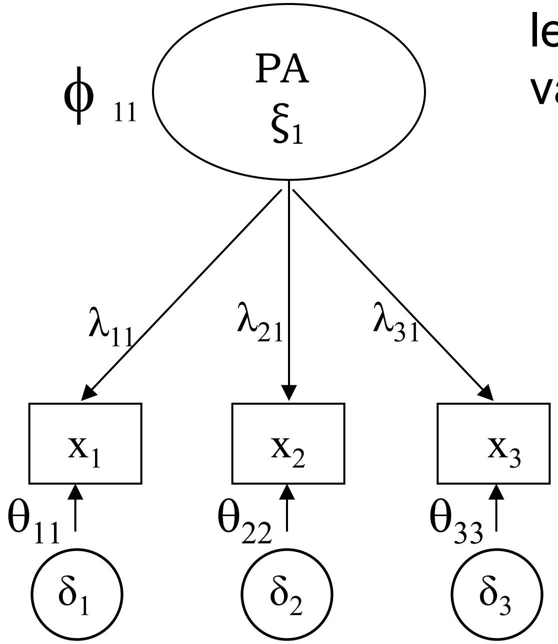
Why is scaling necessary？ the variance parameter depends on the factor loadings：$X = \lambda\phi +$ 。。。

Scaling methods enable the control of the influence of factor loadings in different ways.

Three scaling methods are available！

Scaling the variance parameter

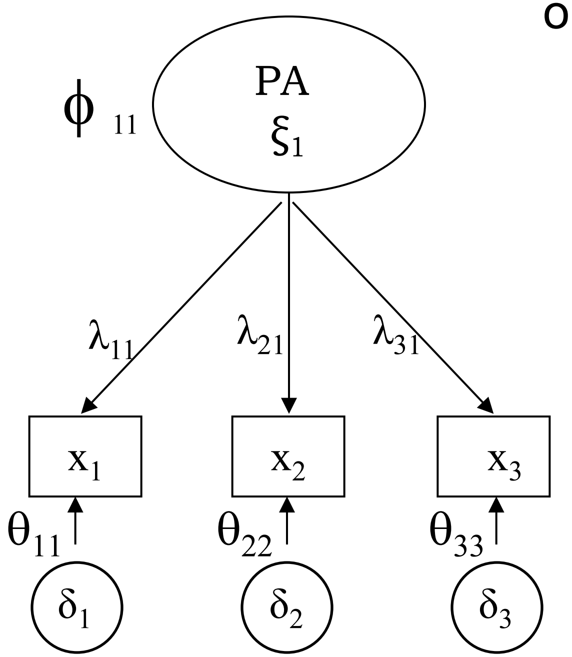
The researcher can be interested in the **estimation of the variance** parameter:



… because comparing estimates for different models can show which model leads to the largest amount of explained variance!

Scaling the variance parameter

The researcher can be interested in the **estimates of the variance** parameters :



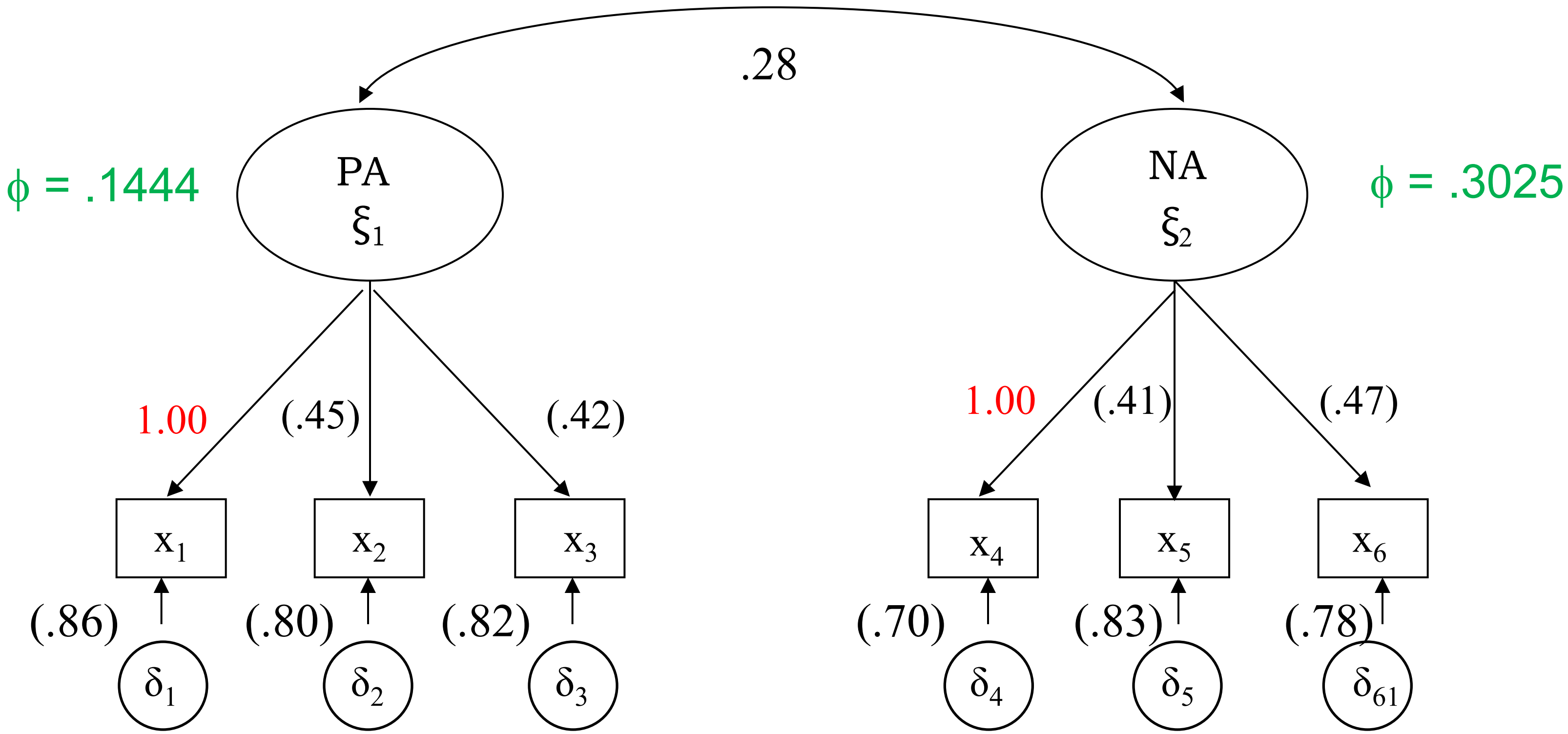
The methods:

1. The *marker variable method*: either λ_{11} or λ_{21} or λ_{31} is fixed to 1 instead of ϕ_{11}.

Disadvantage: the size of ϕ_{11} varies for different lambda.

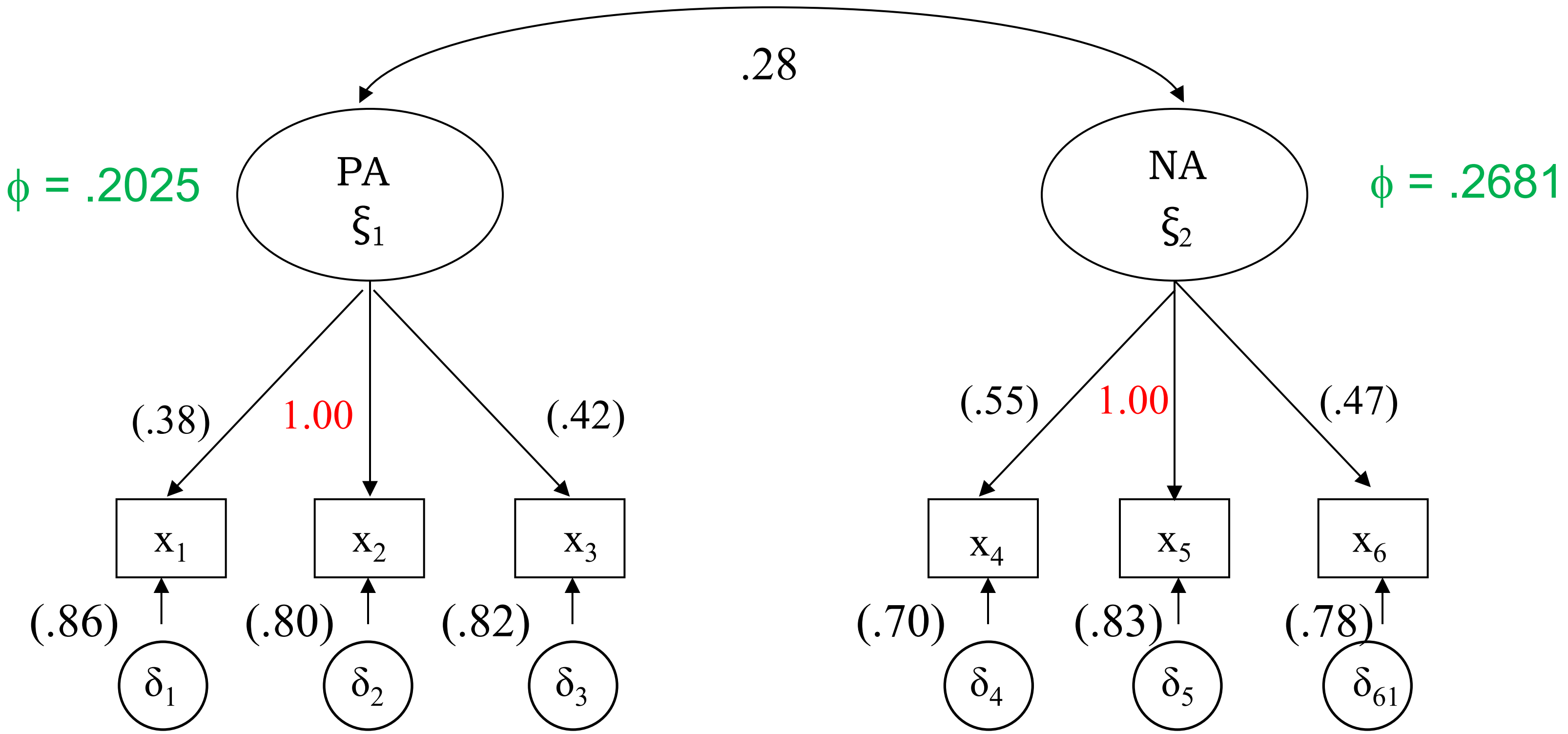
Scaling the variance parameter

Variance estimates according to the marker-variable method



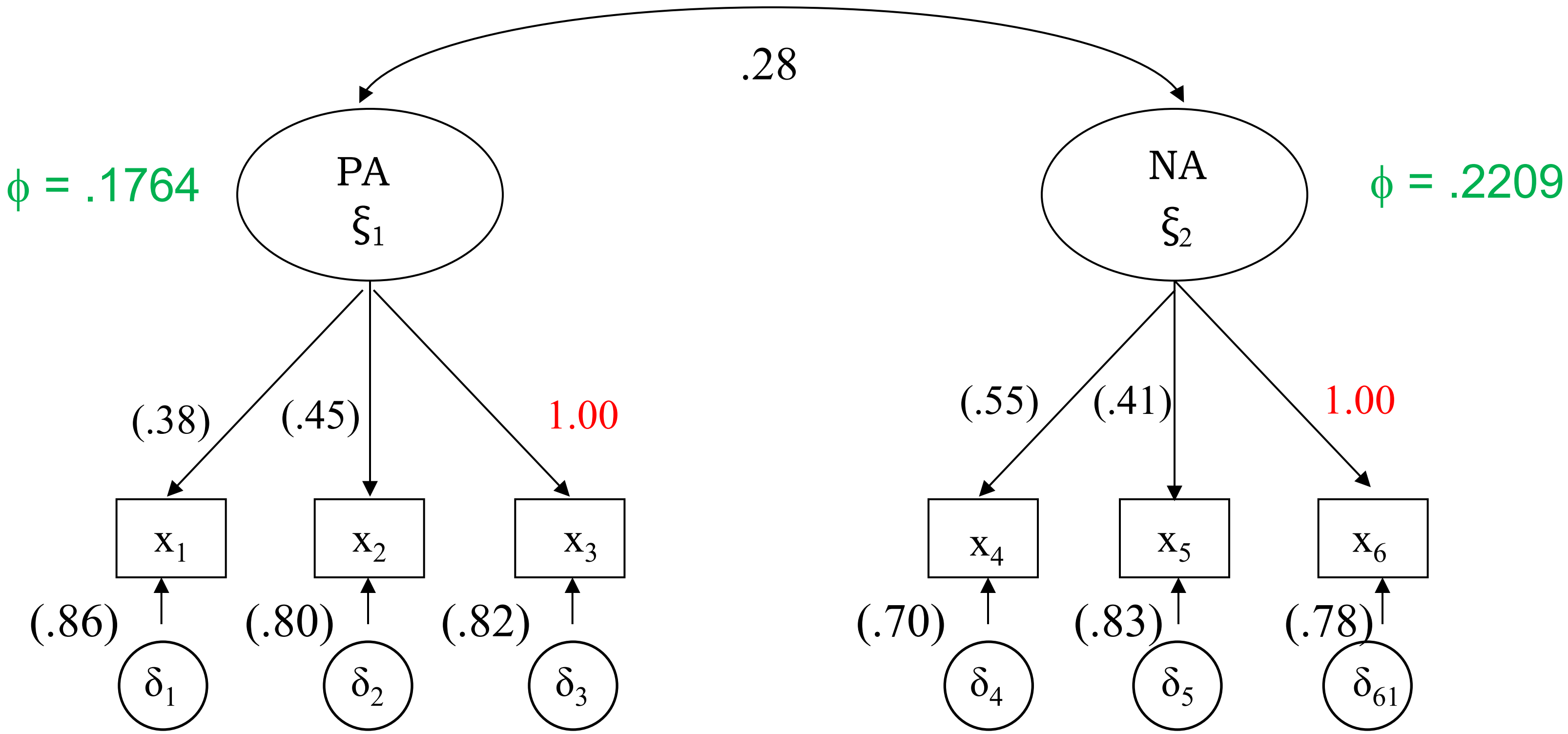
Scaling the variance parameter

Variance estimates according to the marker-variable method



Scaling the variance parameter

Variance estimates according to the marker-variable method

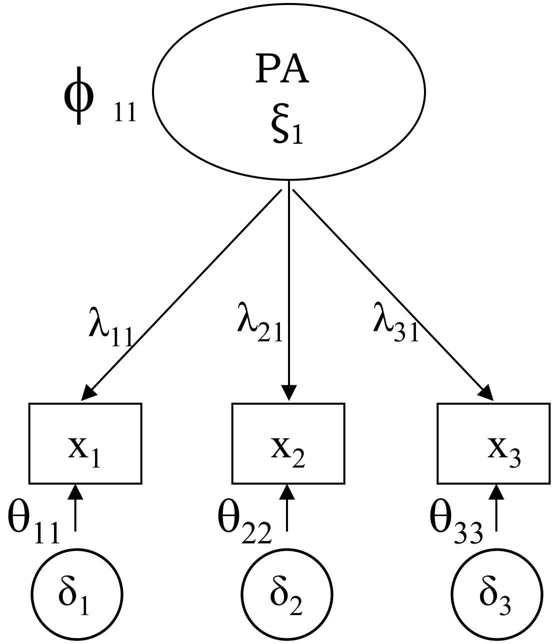


Scaling the variance parameter

The researcher can be interested in the **estimates of the variance** parameters :

The methods:

2. The *<u>reference-group method</u>*: $\phi_{11} = 1$.

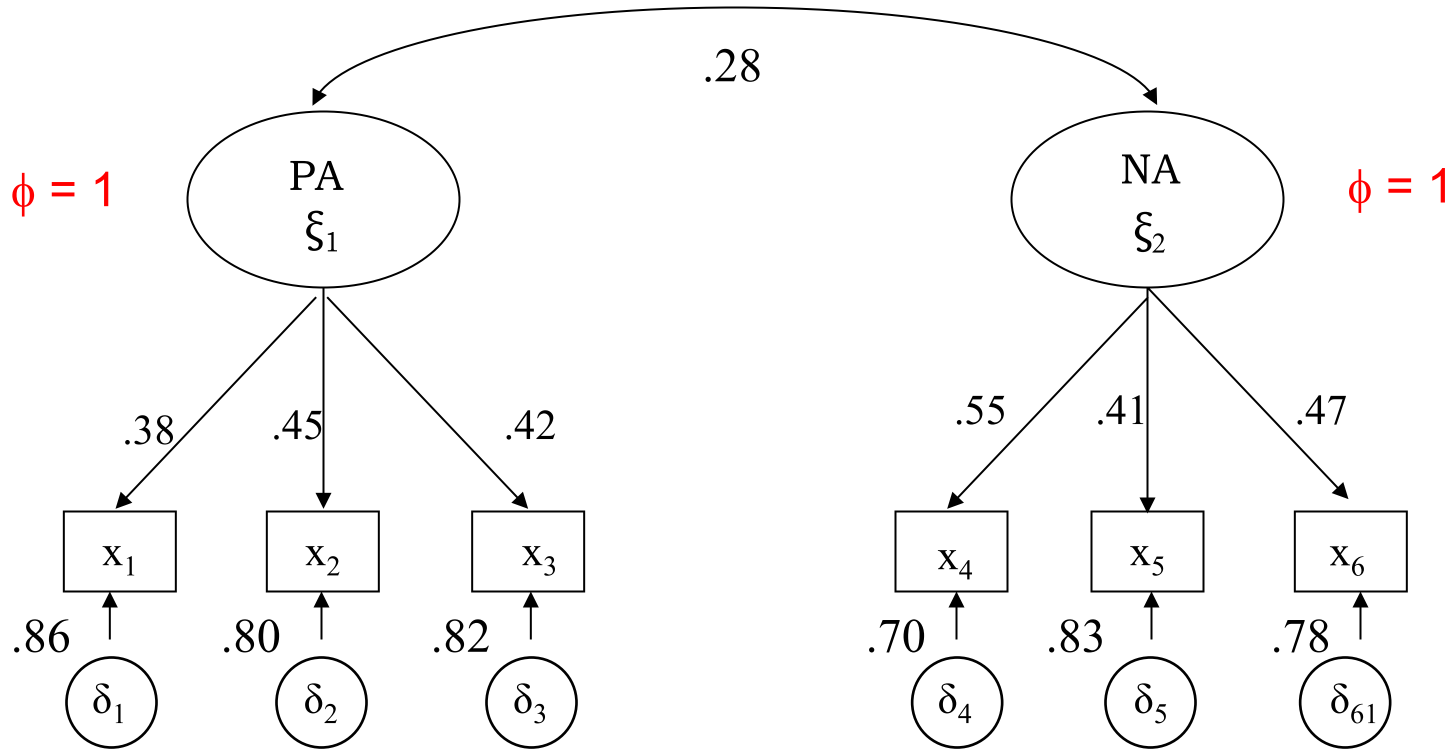


Disadvantage: there is no estimate of the size of ϕ_{11}. It is only useful if there are several groups.

But，the factor variance （ϕ_{11}） can be estimated indirectly as <u>sum of squared factor loadings</u>

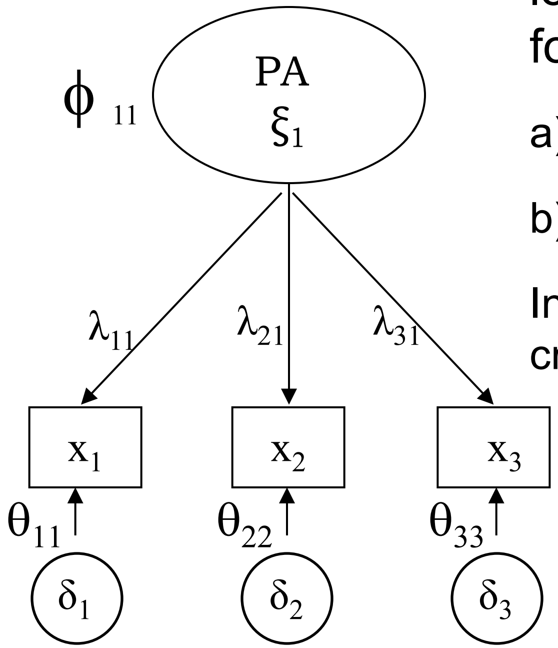
Scaling the variance parameter

Parameter estimates according to the reference group method



Scaling the variance parameter

The researcher can be interested in the **estimates of the variance** parameters :



The methods:

3. The *criterion-based methods*: the factor loadings are standardized in one of the following two ways:

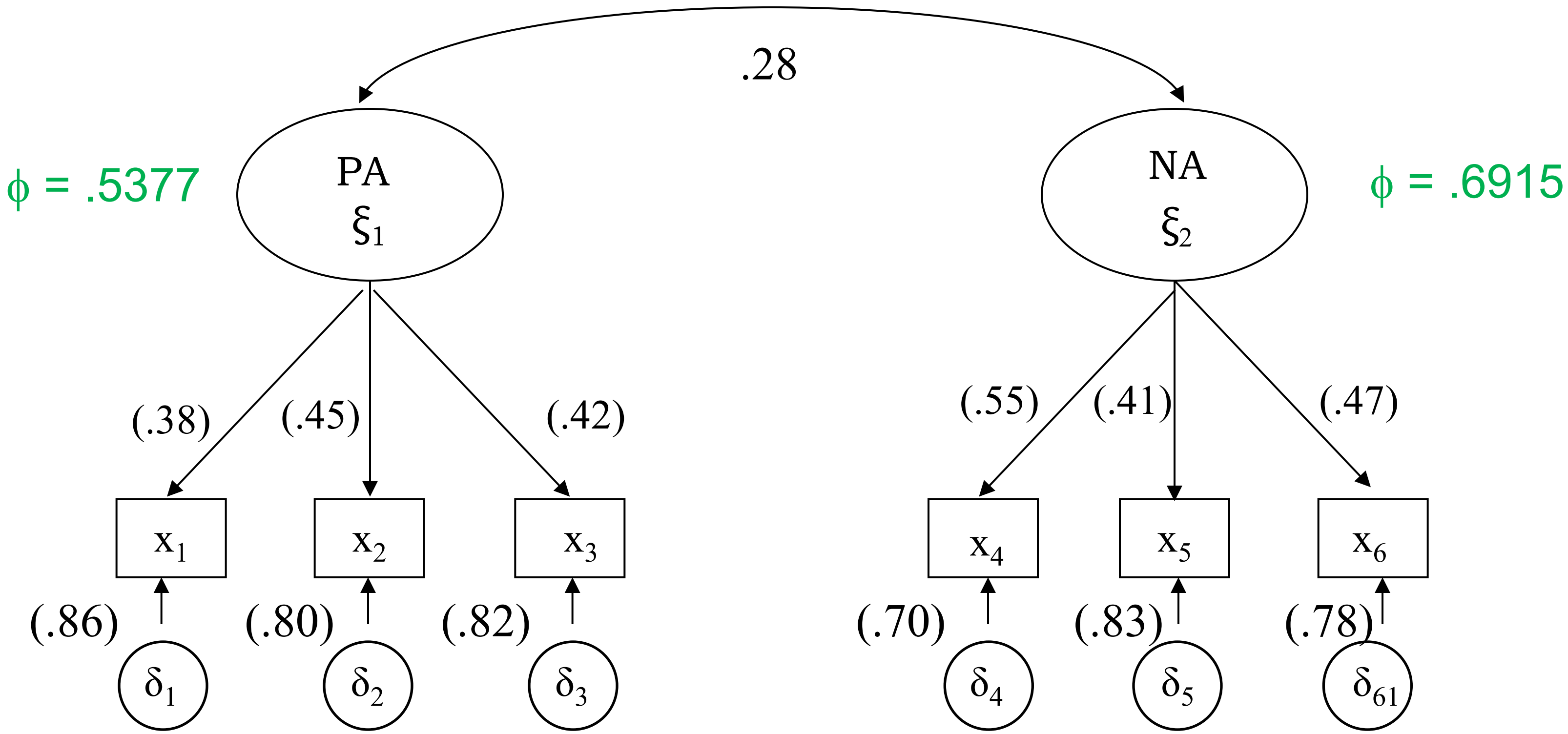
a) criterion = $c\,(\lambda_1 + \ldots + \lambda_p)$ (criterion = p)

b) criterion = $c^2\,(\lambda_1^2 + \ldots + \lambda_p^2)$ (criterion = 1)

In each case c must be selected such that the criterion holds.

Scaling the variance parameter

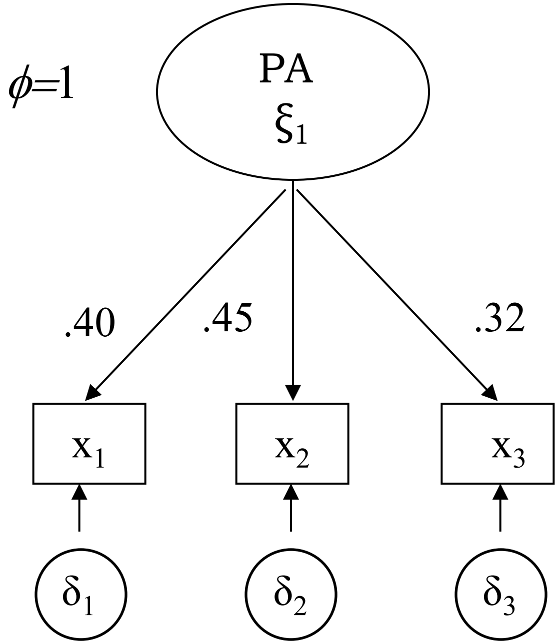
Variance estimates according to the criterion-based method (version 2 (b))



Practice

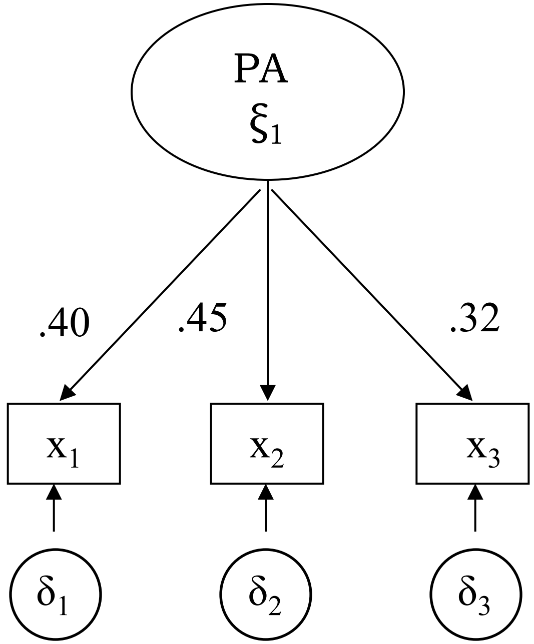
Please, compute the variance of the factor based on the reference-group method（*square the factor loadings and sum them up*）！

Scaling the variance parameter



Determine the factor variance!

Scaling the variance parameter



Determine the ϕ

Proben numbers

- 25
- 33
- 42
- 46
- 52

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- *Evaluation of model fit*
- Model types
- Comparison of EFA and CFA

Evaluation of model fit

- It is checked whether

... the model matix $\mathbf{\Sigma}$ reproduces

... the empirical matrix **S** well enough.

For this purpose the estimated parameters (θ) are used for specifying $\mathbf{\Sigma}$: $\mathbf{\Sigma(}\theta\mathbf{)}$

The aim is ... $\min \mathbf{F}\left[\mathbf{\Sigma}(\boldsymbol{\theta}), \mathbf{S}\right]$

Fitting function

Evaluation of model fit

- It is checked whether ....
  - χ^2 statistic (originates from F(....)
- Use of *fit indices* (some of them ...)
  - Root Mean Square Error of Approximation (RMSEA)
  - Standardized Root Mean Residual (SRMR)
  - Comparative Fit Index (CFI)
  - Non-normed Fit Index (NNFI)

    ....
  - Model difference statistics

Evaluation of model fit

- It is checked whether …
- Use of difference tests:
  - AIC
  - χ^2 difference test
  - **CFI difference test （。01）**
  - **RMSEA difference test （。015）**

Model fit results

Example:

Type of model	χ^2	df	χ^2 / df	RMSEA	SRMR	CFI	NNFI	AIC
One factor	632.0	170	3.71	0.073	0.140	0.952	0.946	712.0
Two factor	476.1	151	3.15	0.065	0.140	0.966	0.957	594.1
Bifactor	210.5	157	1.34	0.026	0.070	0.994	0.993	316.5

Model fit results

Example: model comparison

- CFI Bifactor-One-factor = .042 > .01
- CFI Bifactor-Two-factor = .028 > .01
- RMSEA Bifactor-One-factor = .047 > .015
- RMSEA Bifactor-Two-factor = .039 > .015

Model fit results

Example: model comparison

Δ CFI Bifactor-One-factor = .042 > .01

Δ CFI Bifactor-Two-factor = .028 > .01

Conclusion:
- the bifactor model shows good model fit
- the bifactor model shows the best model fit
- the bifactor model show a substantially better fit than the other models

Practice

Please, determine which model is the to-be-preferred model!

Model fit results

Example:

Type	Case 1	RMSEA	CFI	Case 2	RMSEA	CFI
One factor		0.059	0.942		0.042	0.956
Two factor		0.040	0.955		0.039	0.962

Which model is the to-be-preferred model in cases 1 and 2 ?

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- *Model types*
- Comparison of EFA and CFA

Model types

This section gives an overview on model types that are candidates for the representation of a construct including *two unites*

… for finding the best-fitting model

These model types can be specified with respect to different constructs!

Model types

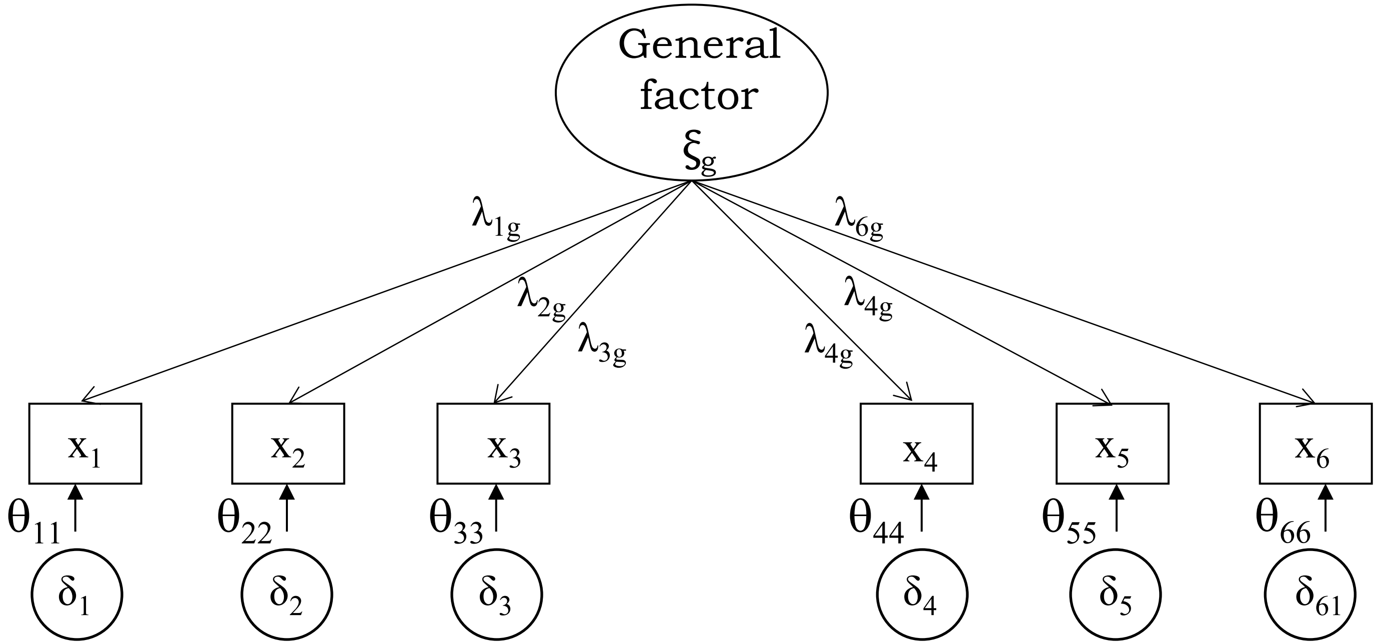
This section gives an overview on model types ....

.... all reasonable structures have to be considered

.... following the simplicity principle the list has to start with the simplest model

Model types

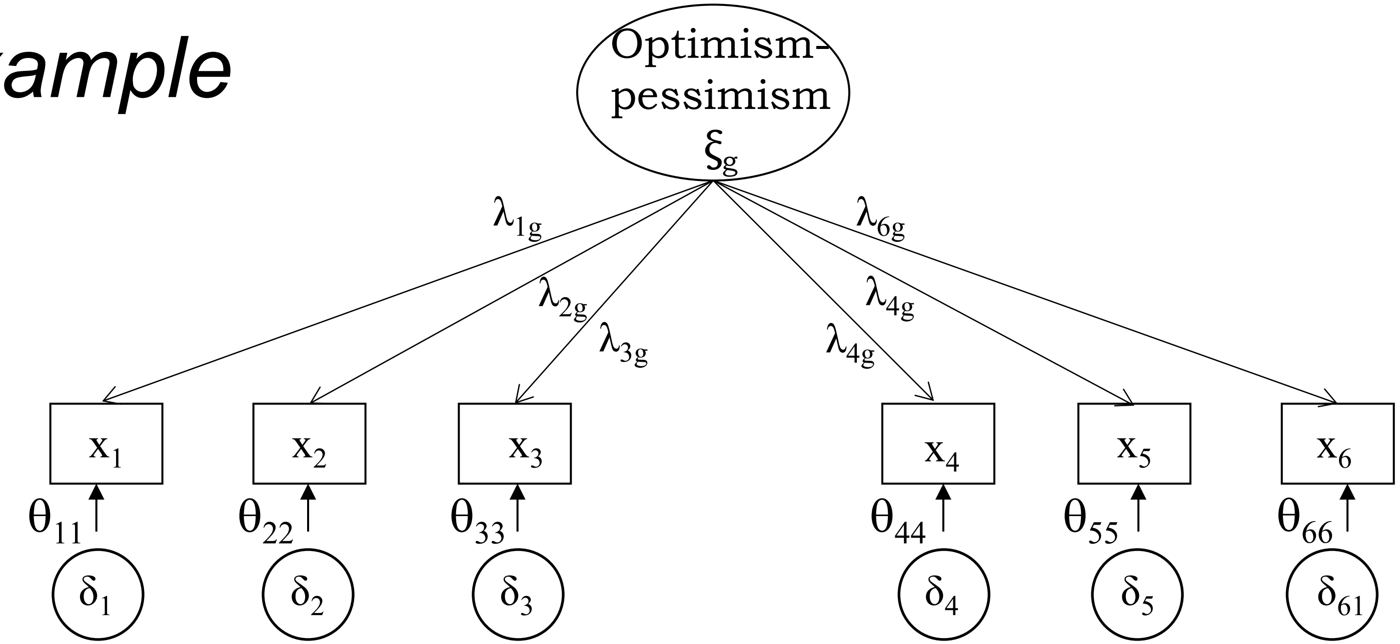
1. The one-factor model: *assuming that both units share the same underlying dimension*



Model types

1. The one-factor model: *assuming that both units share the same underlying dimension*

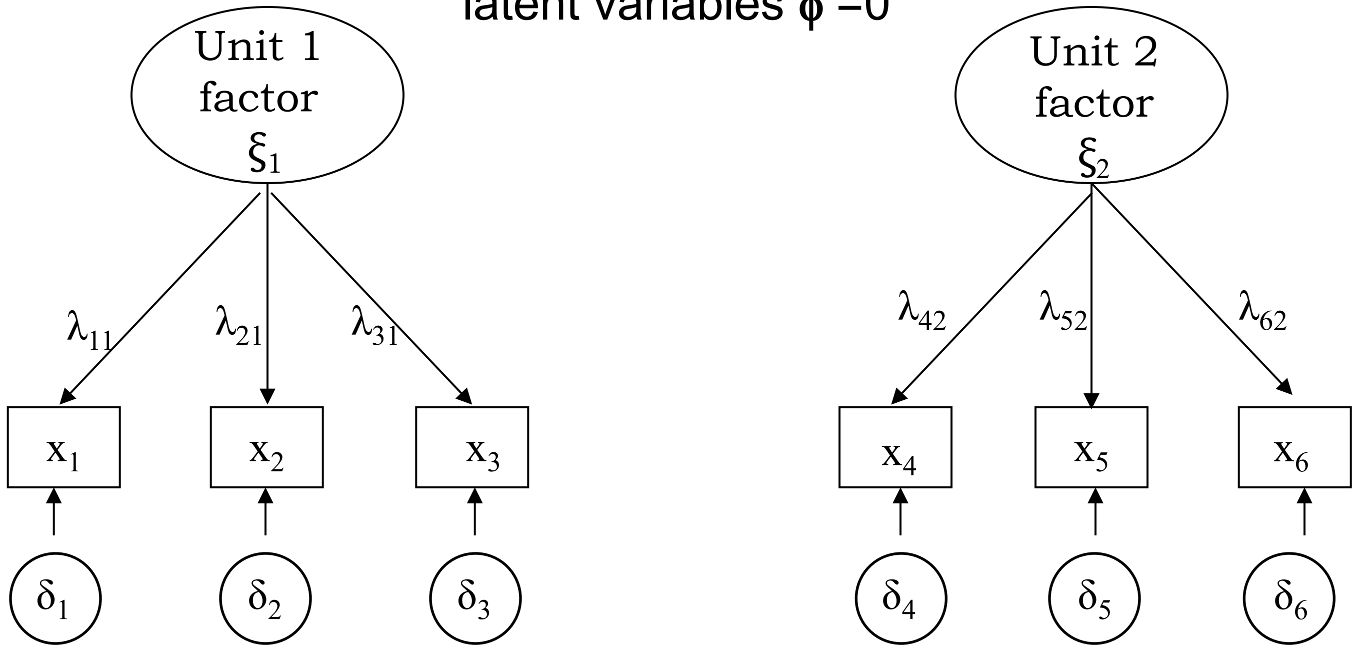
Example



Model types

2.1 The uncorrelated two-factor model: *assuming that the units are unrelated*

No correlation of latent variables ϕ =0

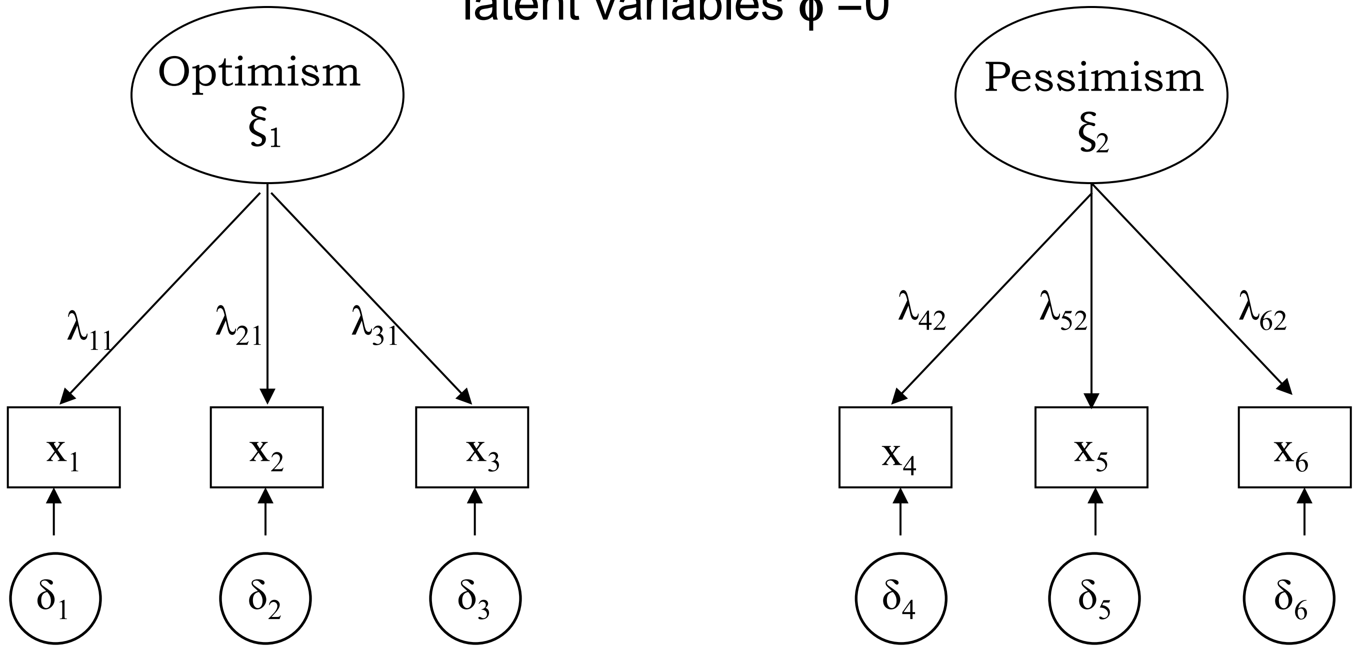


Model types

2.1 The uncorrelated two-factor model: *assuming that the units are unrelated*

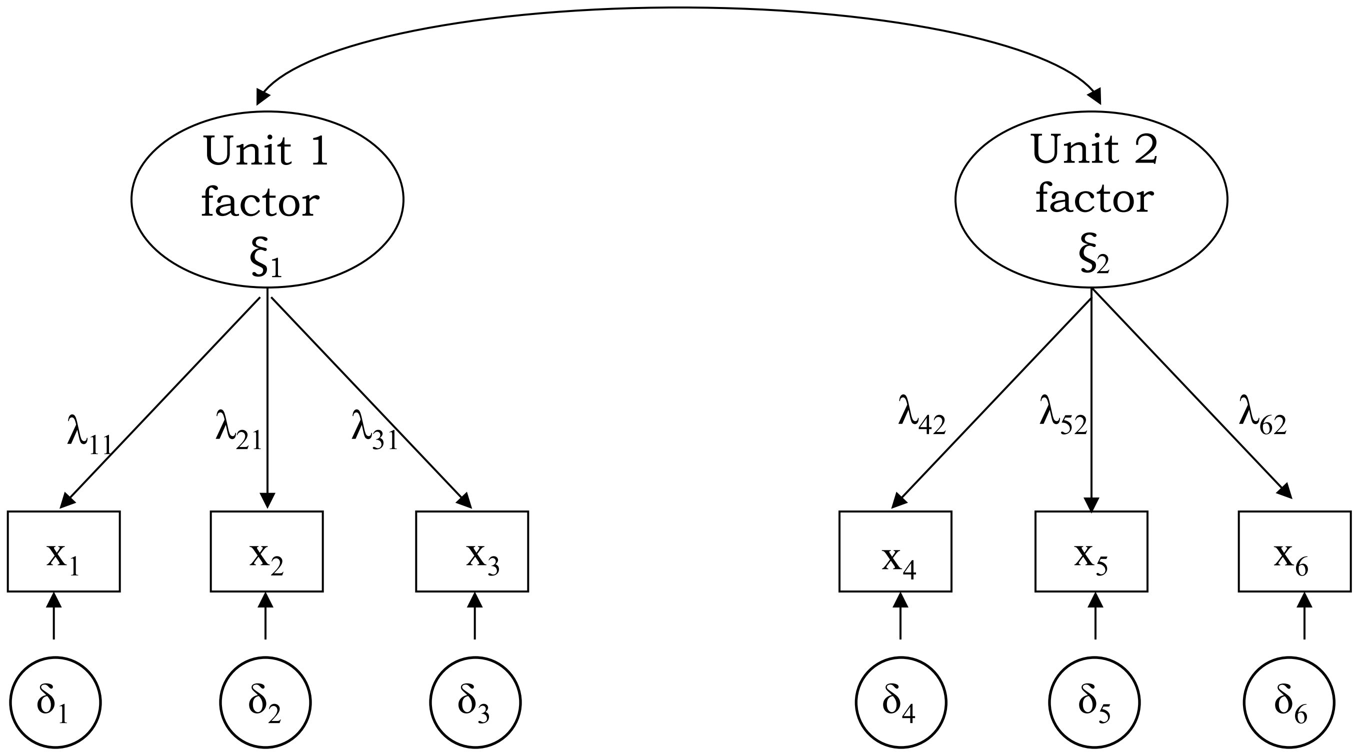
Example

No correlation of
latent variables ϕ =0



Model types

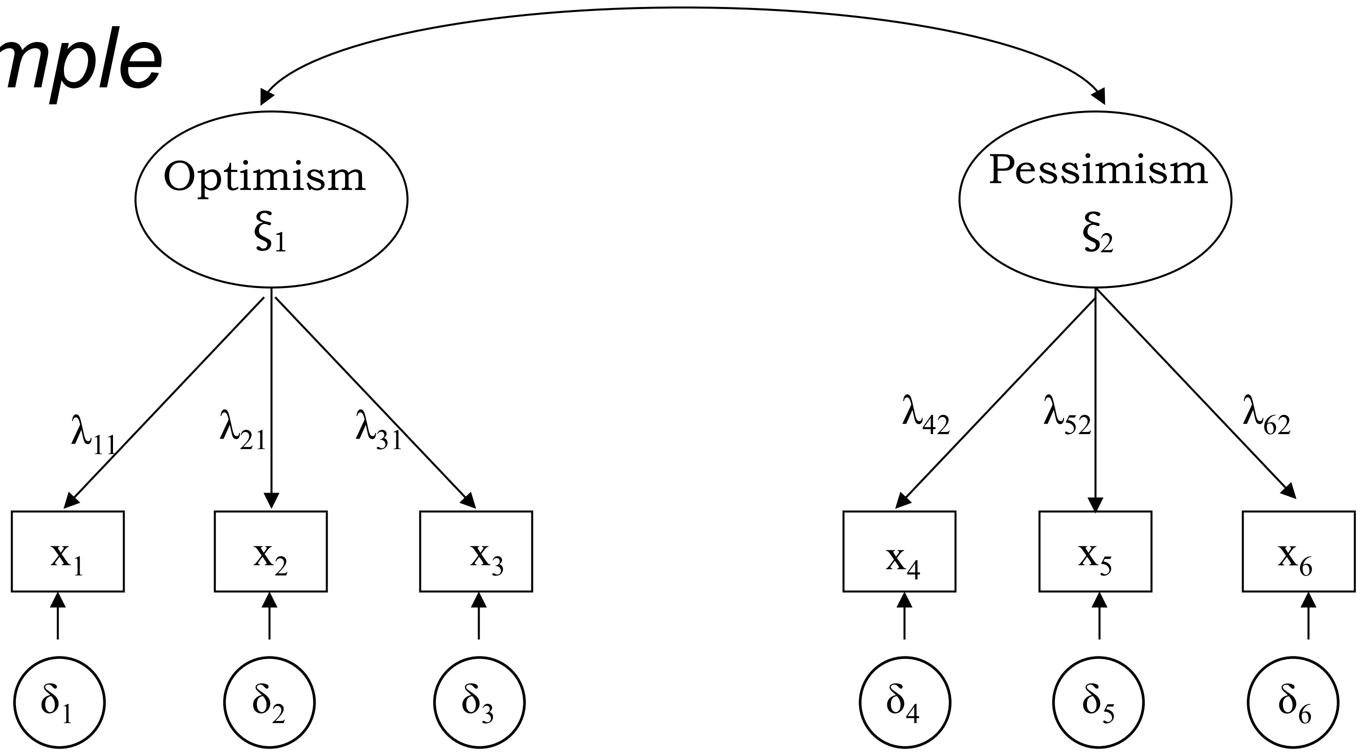
2.2 The correlated two-factor model: *assuming that the units are related*



Model types

2.2 The correlated two-factor model: *assuming that the units are related*

Example



Model types

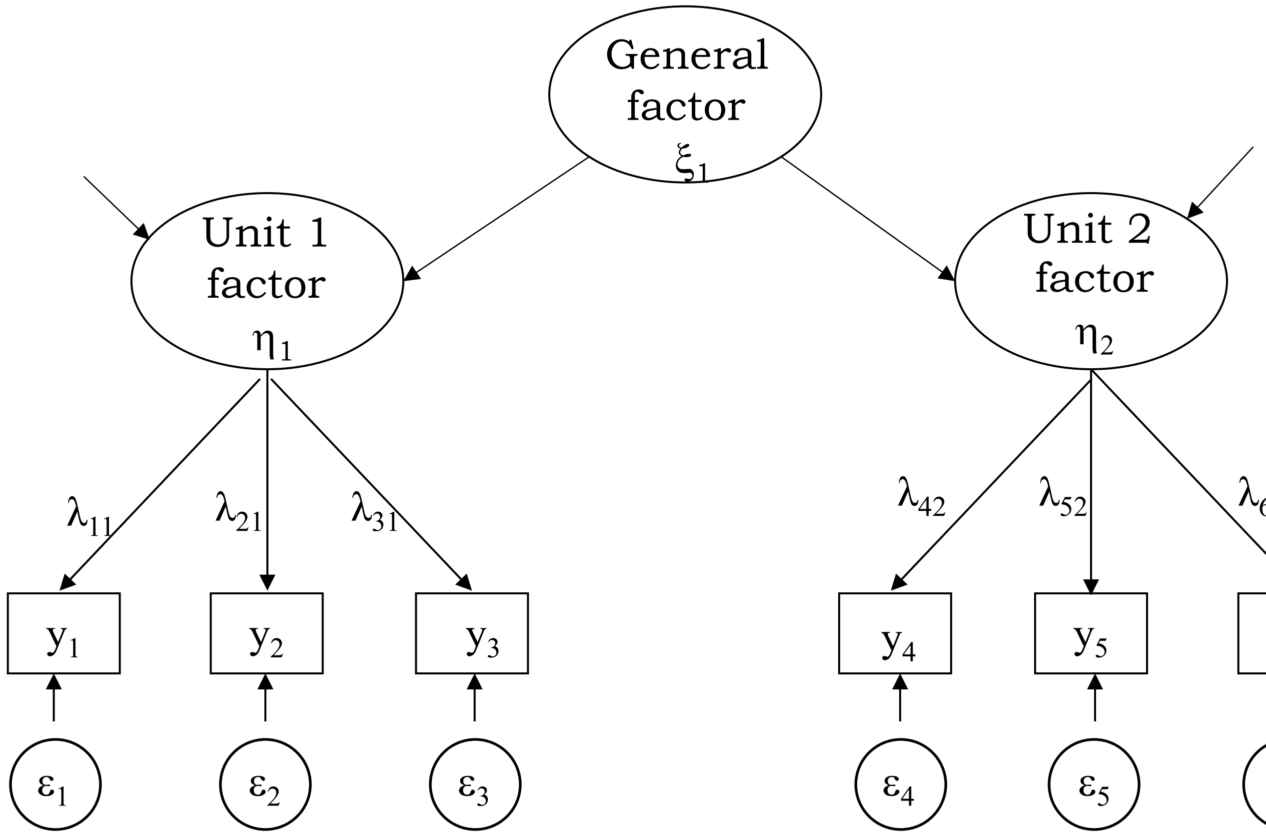
2.2 The correlated two-factor model is usually considered as check for the appropriateness of models 1. and 2.1

… it is (normally) not considered as a good model because the relationship of the latent variables is <u>not explained</u>.

… if it fits better than models 1. and 2.1, this is considered as indication for selecting a more complex model

Model types

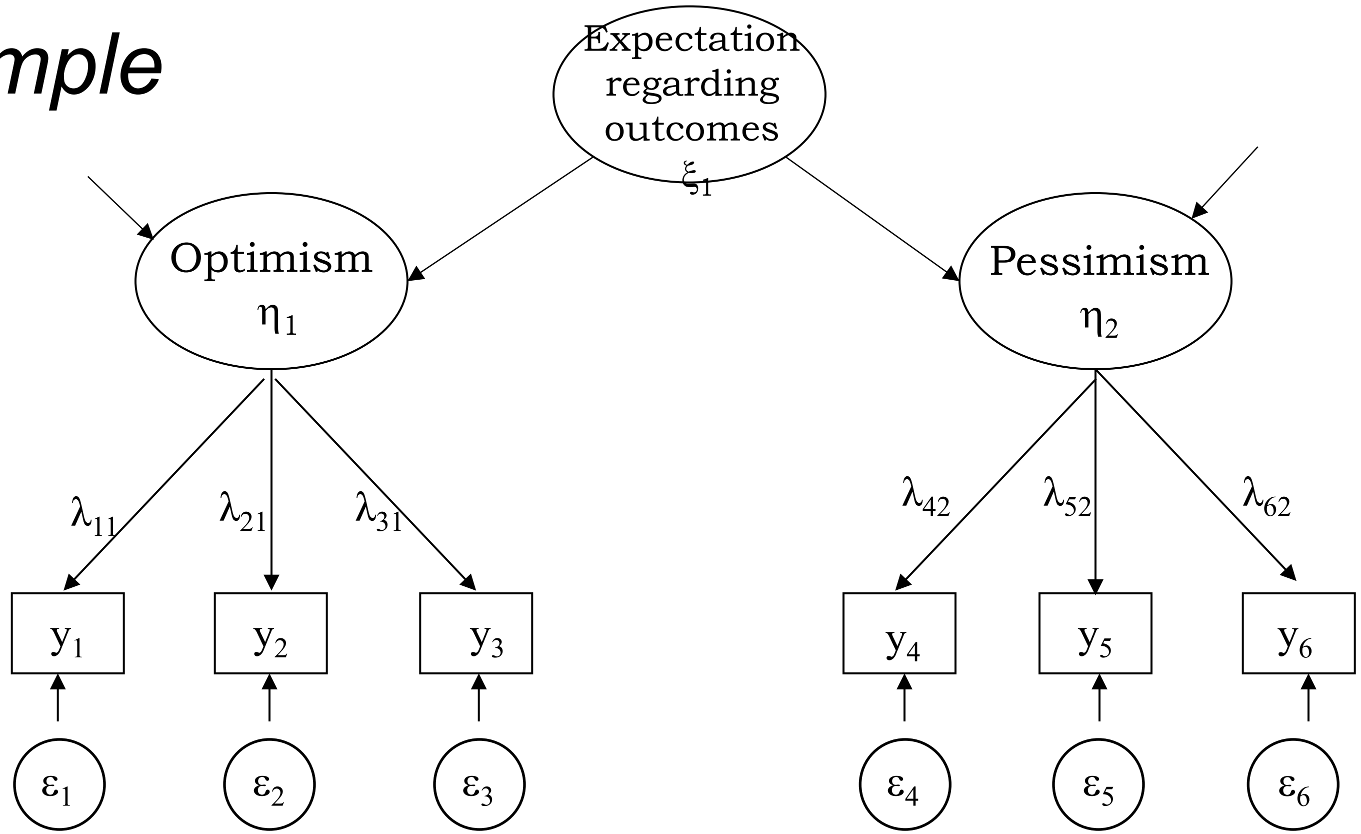
3.1 The hierarchical model



Model types

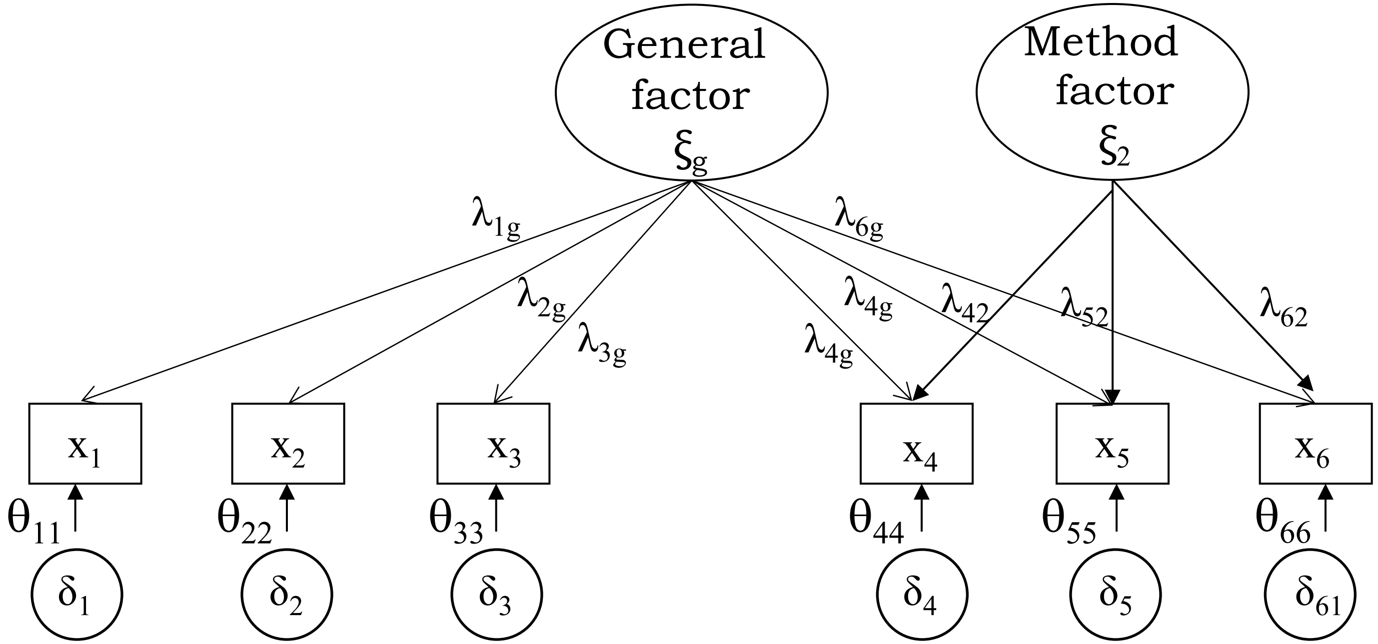
3.1 The hierarchical model

Example



Model types

3.2 Bifactor model with one nested method factor

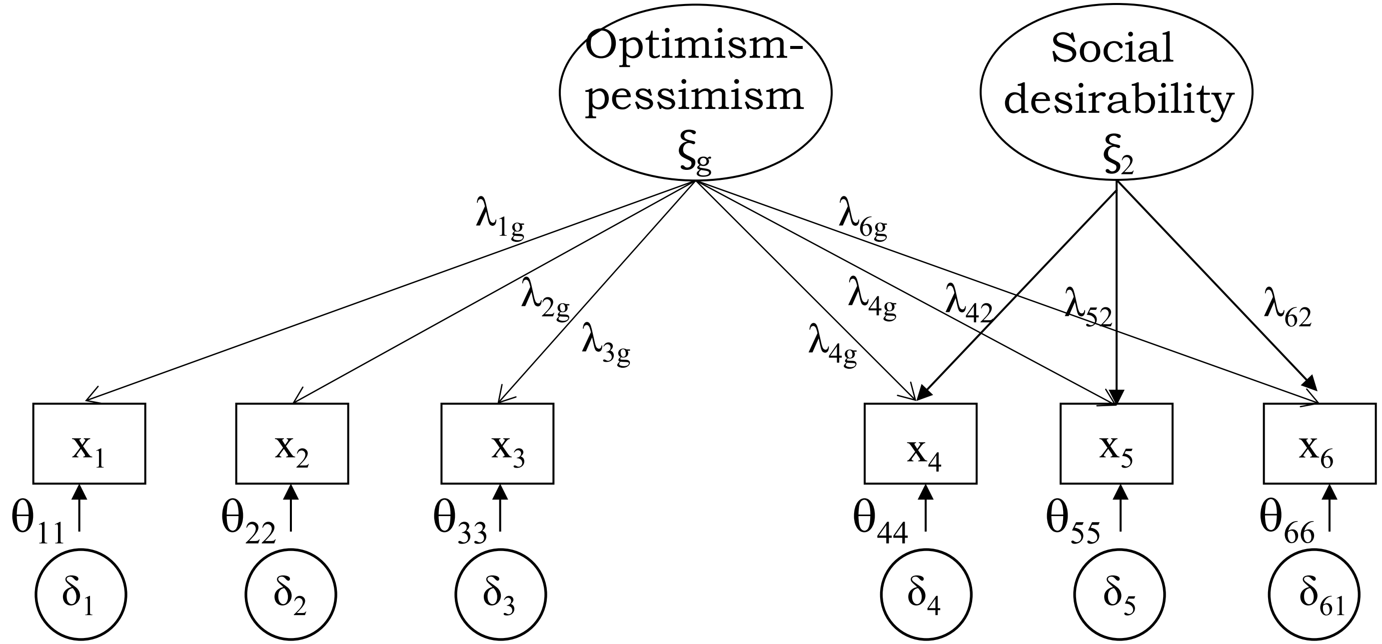


Possible method factors: social desirability, acquiescense, wording, item-position, speed etc.

Model types

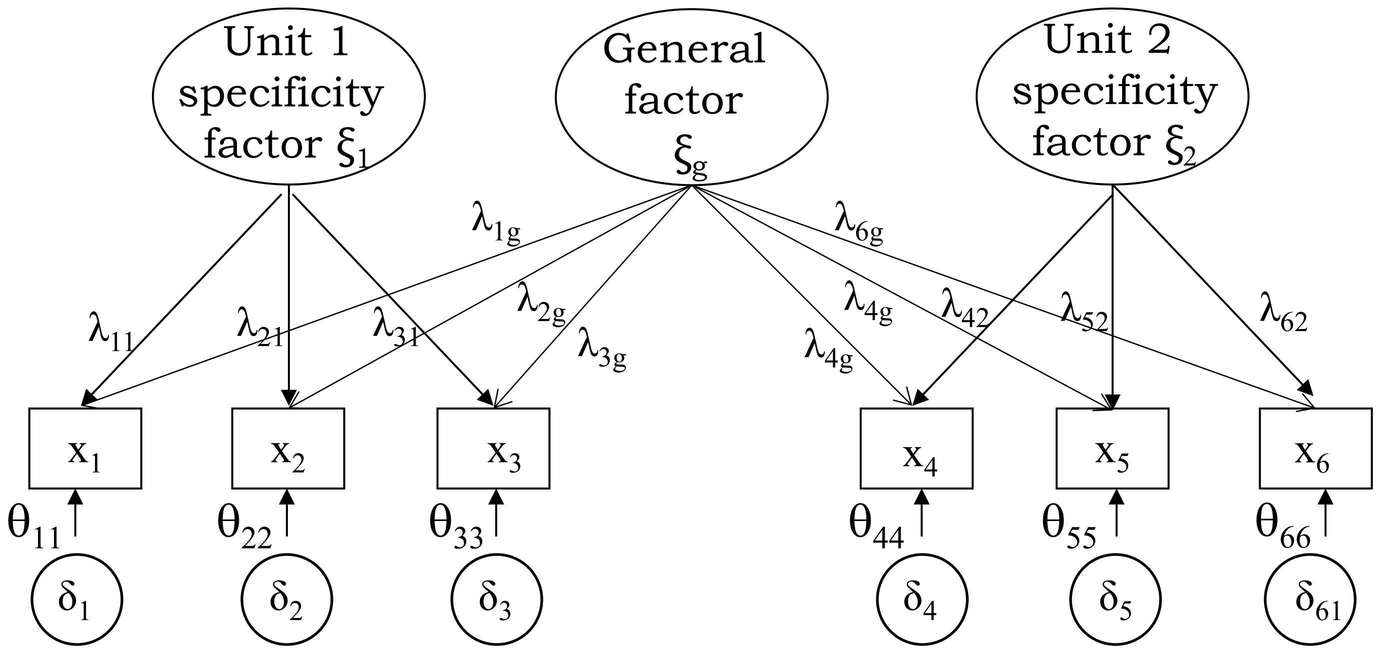
3.2 Bifactor model with one nested method factor

Example



Model types

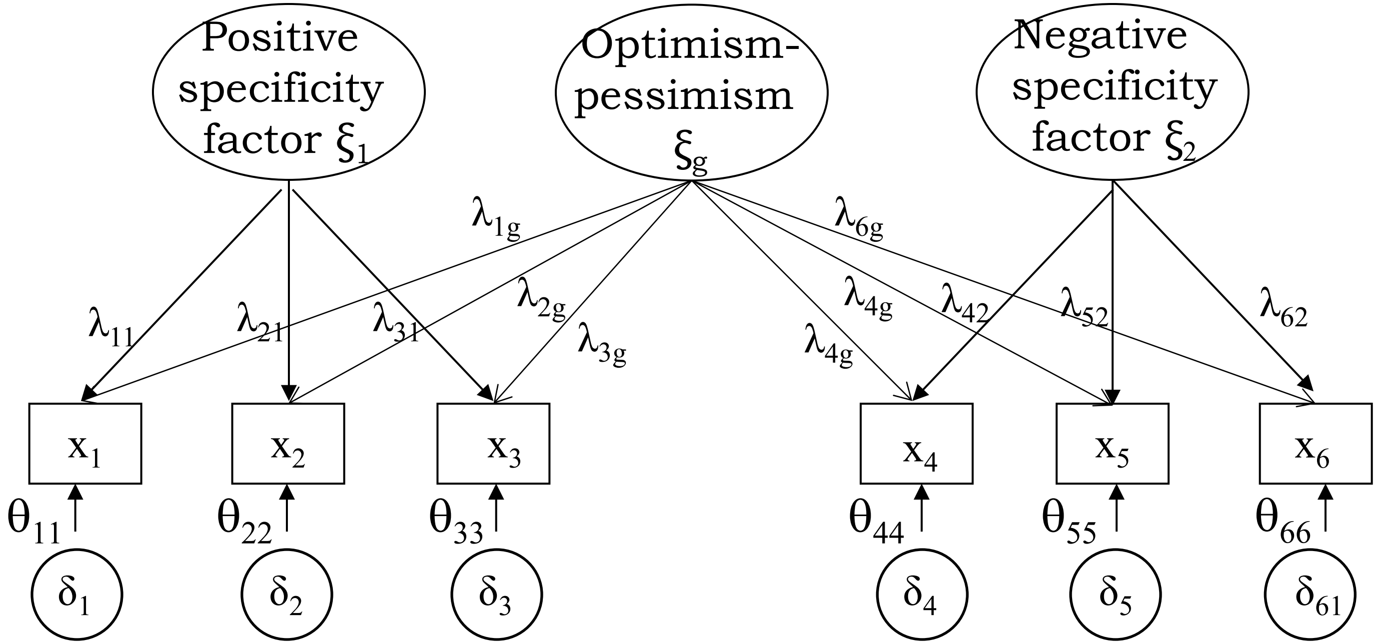
3.3 Extended bifactor model with two nested factors



Model types

3.3 Extended bifactor model with two nested factors

Example



Model types

Table for comparing fit statistics

Model	Structure	$\chi 2$	AIC	CFI $_{\text{difference}}$
1	One-factor	…	…	…
2.1	Uncorrelated two-factor	…	…	…
2.2	Correlated two-factor	…		
3.1	Hierarchical structure	…		
3.2	Bifactor structure	…		
3.3	Extended bifactor structure	…		

Outline

Confirmatory factor analysis

- General remarks
- Formal similarity of EFA and CFA
- Formation of hypotheses
- Construction of path diagram
- Check of identification
- On the research strategy
- Parameter estimation
- Scaling the variance parameter
- Evaluation of model fit
- Model types
- *Comparison of EFA and CFA*

Comparison of EFA and CFA

	EFA	CFA
Type of procedure	Hypothesis-generating procedure	Hypothesis-testing procedure
Data basis	Usually the correlation matrix	Usually the covariance matrix
Number of factors	… during the course of analysis	Fixed a priori
Correlation of factors	Factors are rotated so that all or no factors correlate with each other	It is determined a priori whether individual pairs of factors correlate with each other
Assignment of indicators to factors	All indicators load on all factors	The assignment of the indicators to the factors is based on theory
Comparison of different models	Only one model is considered that emerges in analysis	It is likely that several models are considered, and some of them may be nested
Model quality	Although there are indices of model fit, they are rarely used	There are indices of model fit, and they are regularly used

Summary and brush up:

Confirmatory factor analysis

- General remarks … most important: possible outcomes are known in beforehand
- Formation of hypotheses … know how to state a CFA hypothesis
- Construction of path diagram … make a sketch, name components
- Check of identification … make sure that the model is identified
- On the research strategy … arrange a comparison of models
- Parameter estimation … identify and prepare what needs to be estimated
- Scaling
- Evaluation of model fit … check and interpret the fit indices
- Model types … have an idea of possible alternatives
- Comparison of EFA and CFA … remember that there are similarities but they are not the same

QUESTIONS REGARDING COURSE UNIT 5

- Which type of factor analysis is characterized as hypothesis-testing?
- Which type of matrix is investigated in confirmatory factor analysis?
- What characterizes the research strategy of confirmatory factor analysis?
- What means that there is good model fit?

Literature

Brown, T. A. (2006) *Confirmatory factor analysis.* New York, NJ: The Guilford Press