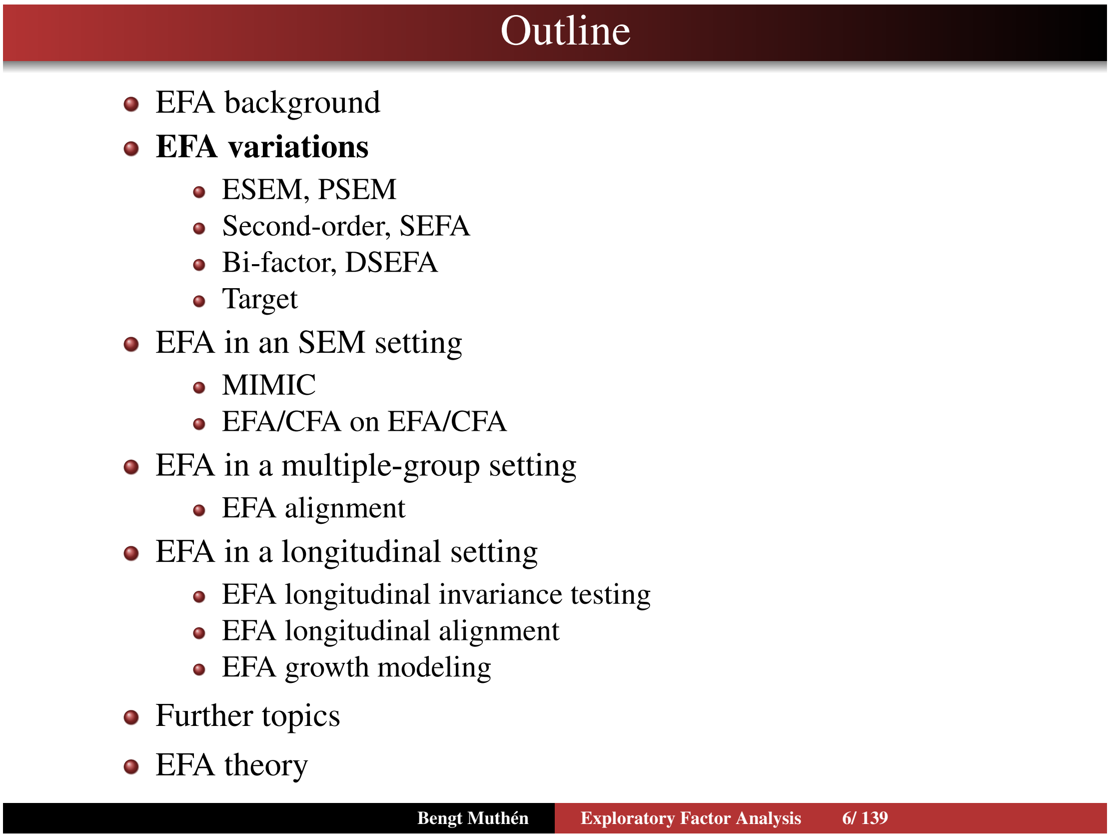
2 EFA variations

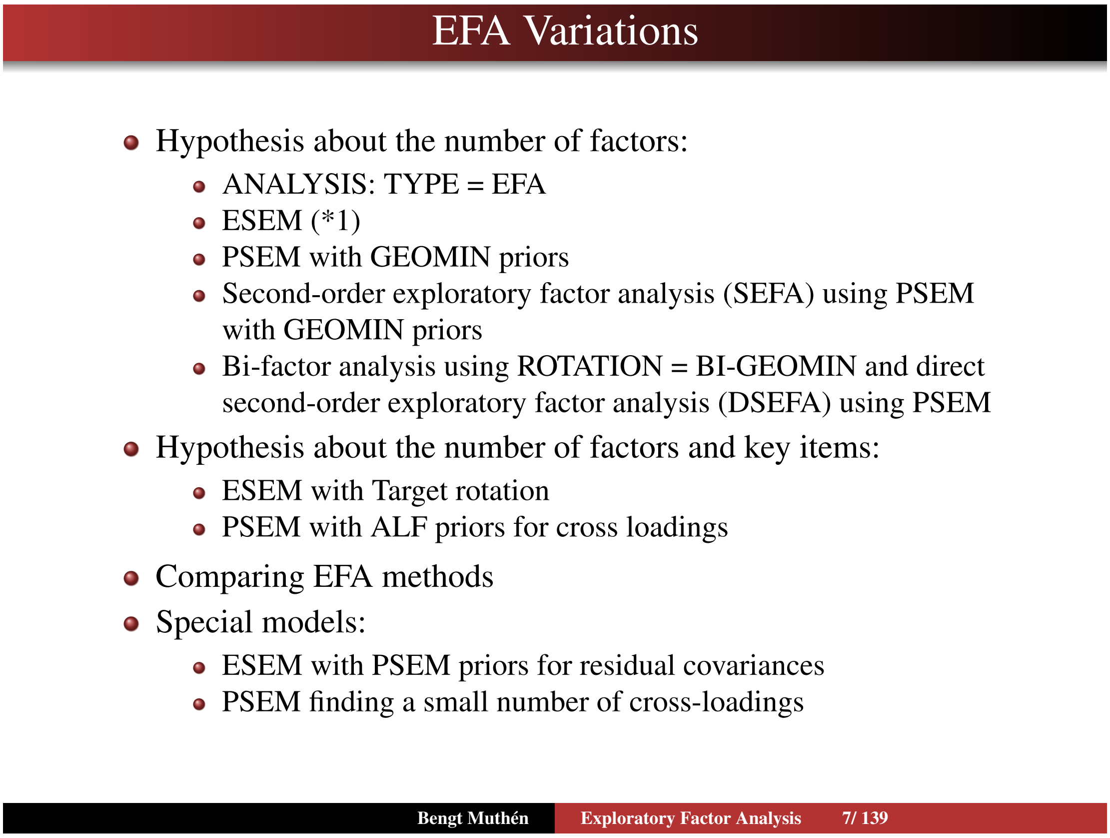


Outline

- EFA background
- **EFA variations**
  - ESEM, PSEM
  - Second-order, SEFA
  - Bi-factor, DSEFA
  - Target
- EFA in an SEM setting
  - MIMIC
  - EFA/CFA on EFA/CFA
- EFA in a multiple-group setting
  - EFA alignment
- EFA in a longitudinal setting
  - EFA longitudinal invariance testing
  - EFA longitudinal alignment
  - EFA growth modeling
- Further topics
- EFA theory

Slide 6 returns to the Outline of the presentation and we turn to the topic of EFA Variations. Here we are going to get into applications of ESEM and PSEM. We will also discuss the new Mplus technique of exploratory second-order factor analysis, abbreviated as SEFA. Exploratory bi-factor analysis will also be covered, including the new Mplus technique of DSEFA - direct effects second-order factor analysis.

2.1 Hypothesis about the number of factors



EFA Variations

- Hypothesis about the number of factors:
    - ANALYSIS: TYPE = EFA
    - ESEM (*1)
    - PSEM with GEOMIN priors
    - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
    - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- Hypothesis about the number of factors and key items:
    - ESEM with Target rotation
    - PSEM with ALF priors for cross loadings
- Comparing EFA methods
- Special models:
    - ESEM with PSEM priors for residual covariances
    - PSEM finding a small number of cross-loadings

The EFA Variations topic is a major part of this presentation and is longer than the other parts. Because of this, slide 7 unpacks this topic, showing its parts. Here, the first two main bullets make a distinction between a hypothesis about the number of factors and hypotheses that also include key items. This distinction is in line with the slide 4 discussion of the early and middle stages of factor analysis.

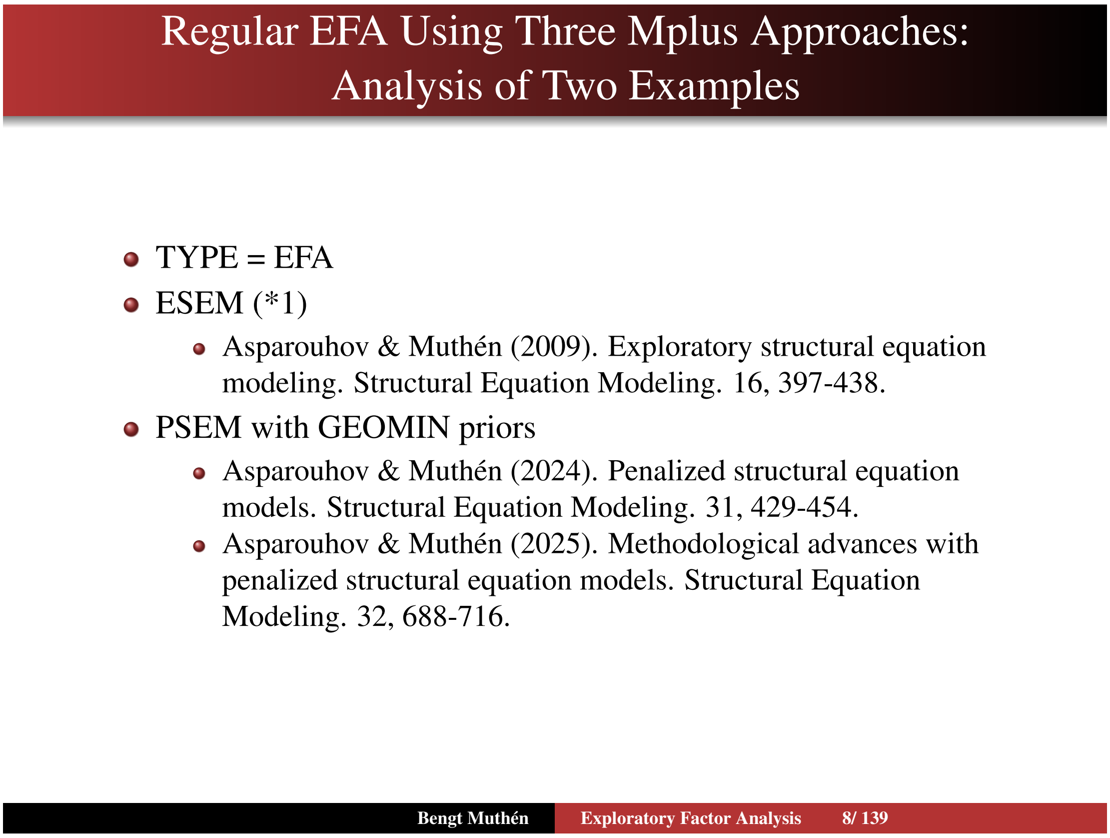
For the main bullet 1, the first 3 sub bullets shows 3 ways that regular EFA can be carried out in Mplus: TYPE = EFA, ESEM, and PSEM GEOMIN. Following this, exploratory versions of second-order and bi-factor analysis using PSEM are discussed.

Main bullet 2 covers cases where key items are specified. Here, two main techniques will be described: ESEM with target rotation, and PSEM ALF priors for cross-loadings.

Main bullet 3 compares EFA methods with a focus on exploratory second-order and bi-factor analyses, but also considers target-oriented approaches.

The EFA Variation section ends with a brief discussion of 2 special models.

2.1.1 Three Regular EFA approaches. H&S example



Regular EFA Using Three Mplus Approaches: Analysis of Two Examples

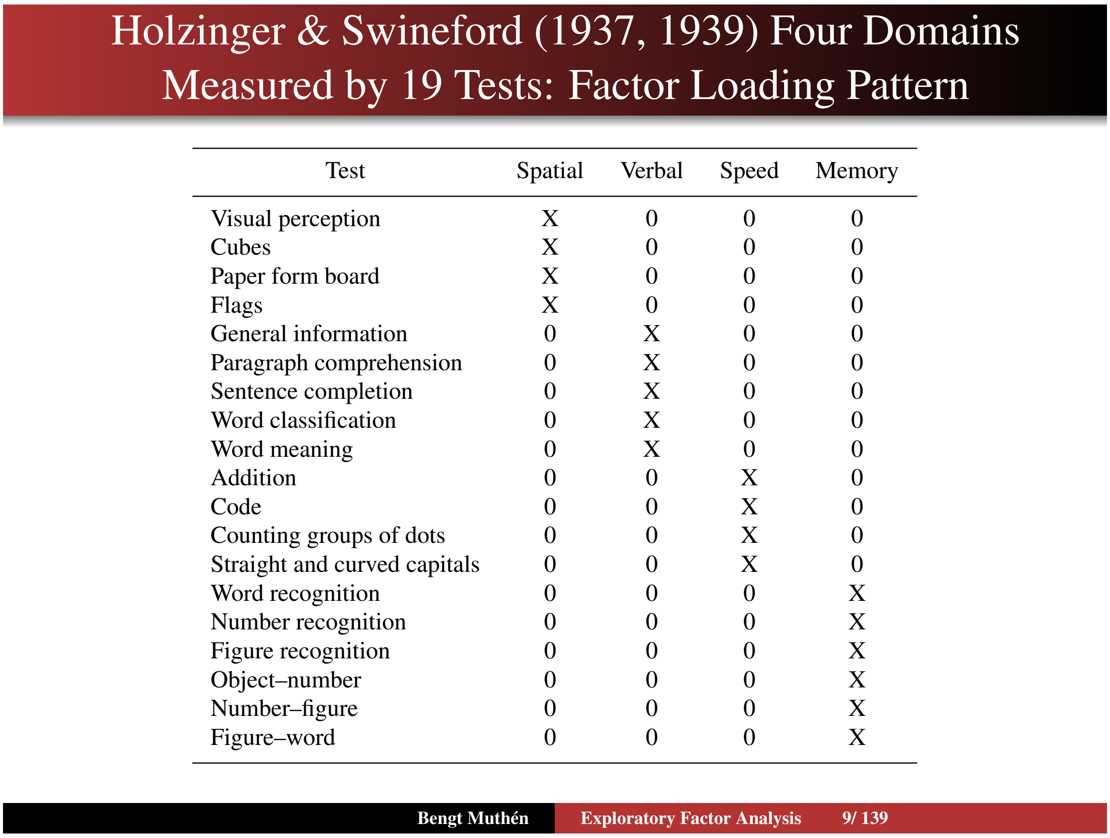
- TYPE = EFA
- ESEM (*1)
  - Asparouhov & Muthén (2009). Exploratory structural equation modeling. Structural Equation Modeling. 16, 397-438.
- PSEM with GEOMIN priors
  - Asparouhov & Muthén (2024). Penalized structural equation models. Structural Equation Modeling. 31, 429-454.
  - Asparouhov & Muthén (2025). Methodological advances with penalized structural equation models. Structural Equation Modeling. 32, 688-716.

We now turn to two examples. Slide 8 shows 3 Mplus approaches to EFA that we will use.

TYPE = EFA allows the analysis of a range of factors in one analysis, computing chi-square difference tests between number of factors.

The ESEM approach gives an equivalent approach for a specific number of factors and also serves as a bridge towards more general EFA settings to be discussed later. The 2009 article is the seminal source of the theory for ESEM.

The PSEM approach gives another equivalent approach using the GEOMIN priors that are the default for the TYPE = EFA approach. As shown in the two articles, this also serves as a bridge towards more general EFA settings to be discussed later.

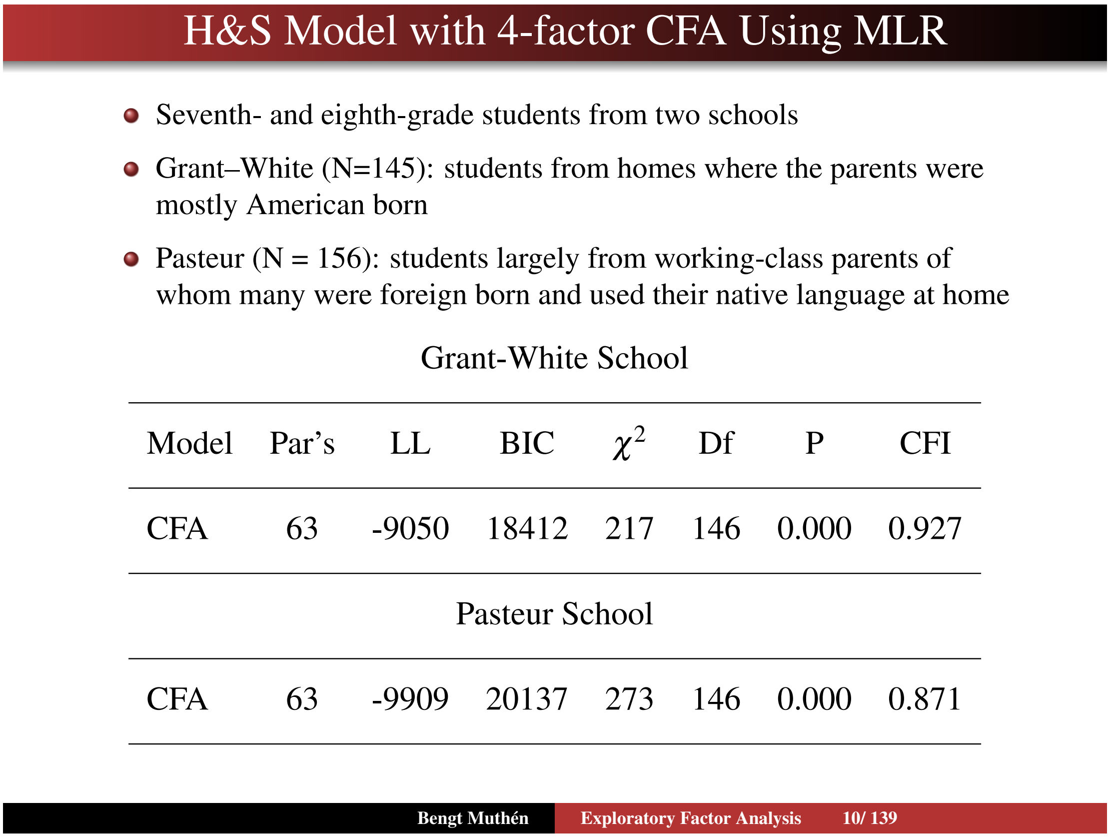


Holzinger & Swineford (1937, 1939) Four Domains Measured by 19 Tests: Factor Loading Pattern

Test	Spatial	Verbal	Speed	Memory
Visual perception	X	0	0	0
Cubes	X	0	0	0
Paper form board	X	0	0	0
Flags	X	0	0	0
General information	0	X	0	0
Paragraph comprehension	0	X	0	0
Sentence completion	0	X	0	0
Word classification	0	X	0	0
Word meaning	0	X	0	0
Addition	0	0	X	0
Code	0	0	X	0
Counting groups of dots	0	0	X	0
Straight and curved capitals	0	0	X	0
Word recognition	0	0	0	X
Number recognition	0	0	0	X
Figure recognition	0	0	0	X
Object–number	0	0	0	X
Number–figure	0	0	0	X
Figure–word	0	0	0	X

On slide 9 we have our first example. This is the classic dataset used in the 1937 and 1939 Holzinger-Swineford papers given in the references. The figure shows a loading matrix for 19 tests with X marking hypothesized large loadings and 0 elsewhere. Four factors are measured by these tests: Spatial, Verbal, Speed, and Memory. Later on we will add 5 more tests and look at the bi-factor model of primary interest to Holzinger and Swineford.

In terms of analysis stages discussed on slide 4, having this hypothesized loading pattern corresponds to the late analysis stage of CFA but here we will primarily focus on EFA approaches. The dataset is available on the web paper page.



H&S Model with 4-factor CFA Using MLR

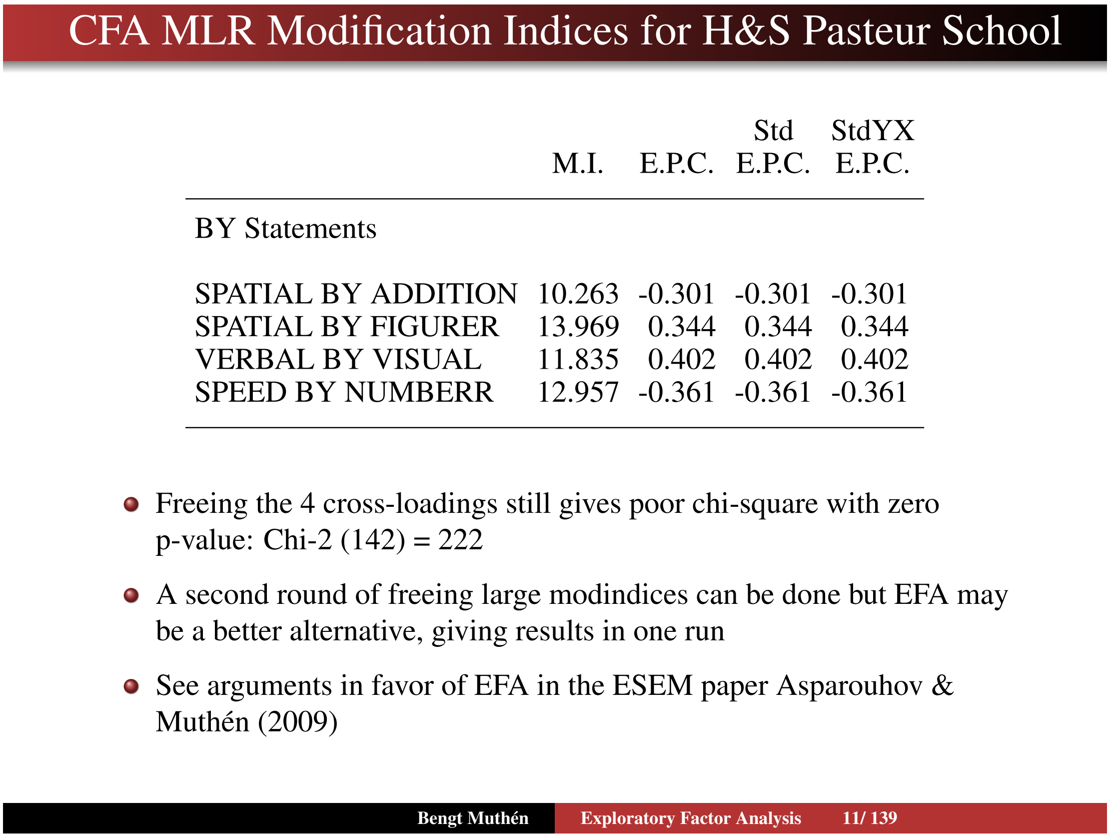
- Seventh- and eighth-grade students from two schools
- Grant–White (N=145): students from homes where the parents were mostly American born
- Pasteur (N = 156): students largely from working-class parents of whom many were foreign born and used their native language at home

Grant-White School

Model	Par's	LL	BIC	χ^2	Df	P	CFI
CFA	63	-9050	18412	217	146	0.000	0.927
Pasteur School							
CFA	63	-9909	20137	273	146	0.000	0.871

On slide 10, we will first show results from a 4-factor CFA using the loading pattern of the previous slide.

The dataset consists of 7th and 8th grade students from two schools, Grant-White and Pasteur. Each has a sample of about 150 students. As the table shows, both schools get a poorly fitting model as judged by chi-square with zero p-value. Given the moderate sample size, the poor fit cannot be ascribed to rejection due to a large sample. Also supporting the rejection, the MLR estimator is used which reduces the sensitivity of the chi-square test to nonnormality in the factor indicators.



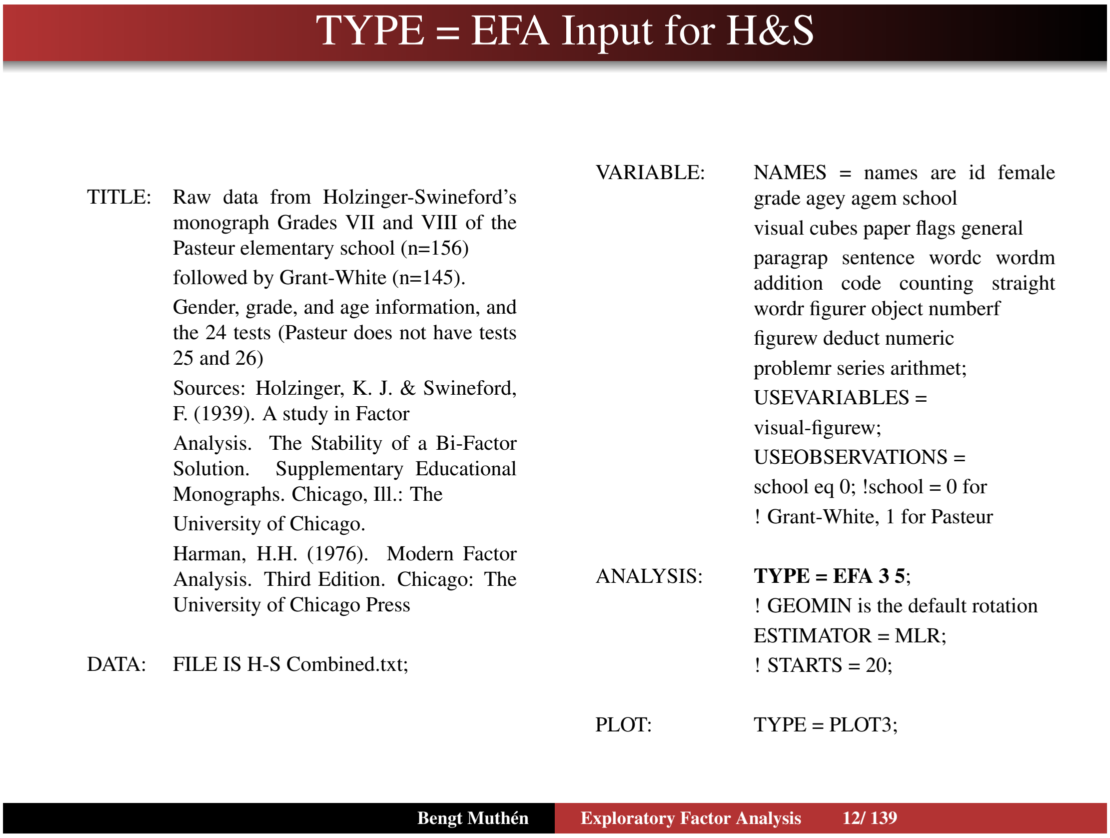
CFA MLR Modification Indices for H&S Pasteur School

			Std	StdYX
	M.I.	E.P.C.	E.P.C.	E.P.C.
BY Statements				
SPATIAL BY ADDITION	10.263	-0.301	-0.301	-0.301
SPATIAL BY FIGURER	13.969	0.344	0.344	0.344
VERBAL BY VISUAL	11.835	0.402	0.402	0.402
SPEED BY NUMBERR	12.957	-0.361	-0.361	-0.361

- Freeing the 4 cross-loadings still gives poor chi-square with zero p-value: Chi-2 (142) = 222
- A second round of freeing large modindices can be done but EFA may be a better alternative, giving results in one run
- See arguments in favor of EFA in the ESEM paper Asparouhov & Muthén (2009)

Slide 11 shows the key modification indices for the CFA run corresponding to unexpected cross loadings. The expected parameter change values for these are rather large in the standardized scale so these are substantively important misspecifications.

As the first bullet says, however, freeing these 4 cross loadings still results in poor fit. Instead of going further as guided by modification indices, EFA is a better approach in line with arguments mentioned on slide 4 in the ESEM paper quote of Brown.

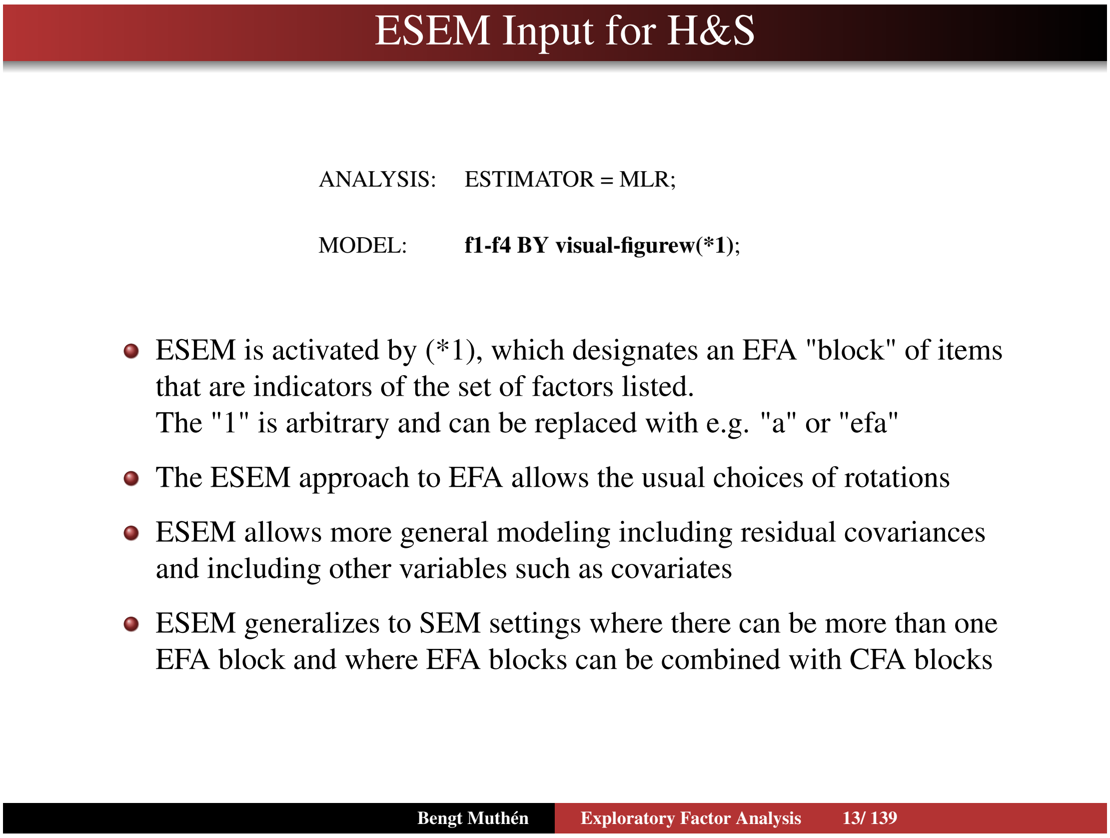


TYPE = EFA Input for H&S

```
TITLE:     Raw data from Holzinger-Swineford's
           monograph Grades VII and VIII of the
           Pasteur elementary school (n=156)
           followed by Grant-White (n=145).
           Gender, grade, and age information, and
           the 24 tests (Pasteur does not have tests
           25 and 26)
           Sources: Holzinger, K. J. & Swineford,
           F. (1939). A study in Factor
           Analysis. The Stability of a Bi-Factor
           Solution. Supplementary Educational
           Monographs. Chicago, Ill.: The
           University of Chicago.
           Harman, H.H. (1976). Modern Factor
           Analysis. Third Edition. Chicago: The
           University of Chicago Press

DATA:      FILE IS H-S Combined.txt;

VARIABLE:  NAMES = names are id female
           grade agey agem school
           visual cubes paper flags general
           paragrap sentence wordc wordm
           addition code counting straight
           wordr figurer object numberf
           figurew deduct numeric
           problemr series arithmet;
           USEVARIABLES =
           visual-figurew;
           USEOBSERVATIONS =
           school eq 0; !school = 0 for
           ! Grant-White, 1 for Pasteur

ANALYSIS:  TYPE = EFA 3 5;
           ! GEOMIN is the default rotation
           ESTIMATOR = MLR;
           ! STARTS = 20;

PLOT:      TYPE = PLOT3;
```

Slide 12 shows the input for EFA where in the ANALYSIS command, EFA with 3, 4, and 5 factors is requested. GEOMIN is the default rotation.

For some special situations, it may be appropriate to add several random starting values using the STARTS option. Here, 20 such sets of starting values are used.



ESEM Input for H&S

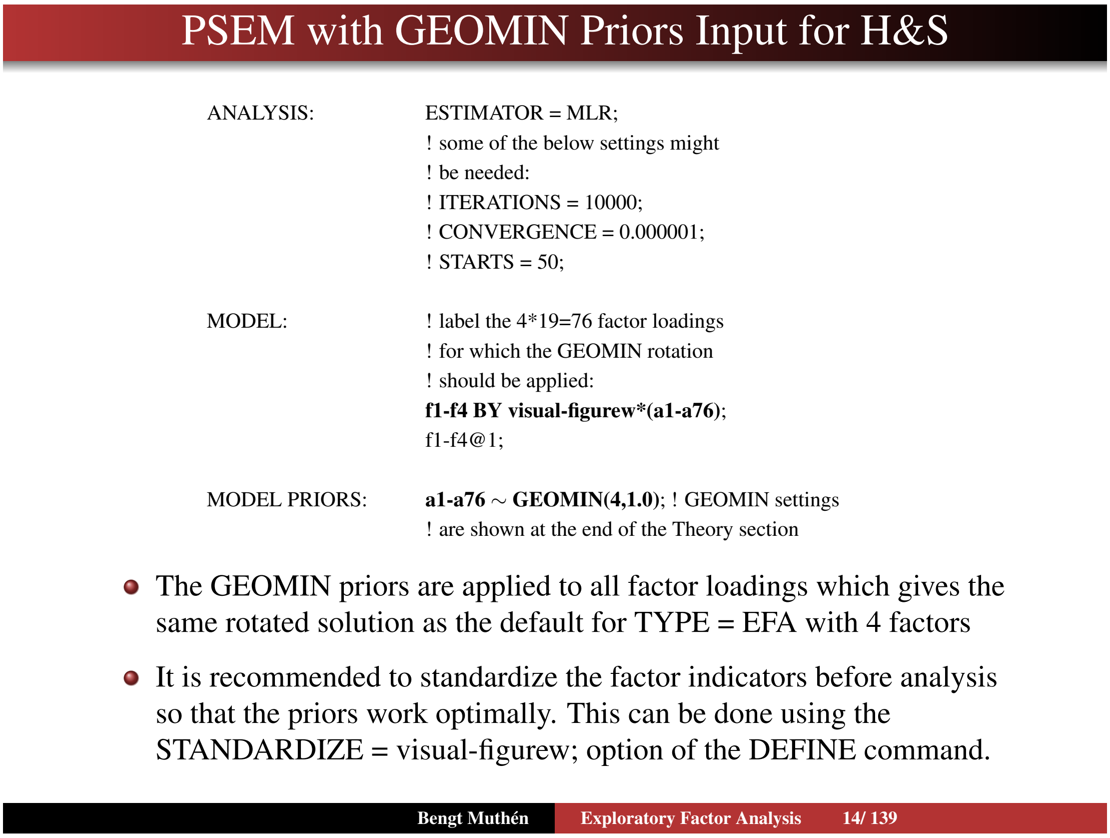
```
ANALYSIS:   ESTIMATOR = MLR;

MODEL:      f1-f4 BY visual-figurew(*1);
```

- ESEM is activated by (*1), which designates an EFA "block" of items that are indicators of the set of factors listed.
  The "1" is arbitrary and can be replaced with e.g. "a" or "efa"
- The ESEM approach to EFA allows the usual choices of rotations
- ESEM allows more general modeling including residual covariances and including other variables such as covariates
- ESEM generalizes to SEM settings where there can be more than one EFA block and where EFA blocks can be combined with CFA blocks

Slide 13 shows the ESEM approach for 4 factors. ESEM is activated by (*1), which designates an EFA “block” of items that are indicators of the set of factors listed, such as f1-f4 BY visual-figurew. The “1” is arbitrary and can be replaced with e.g. “a” or “efa”.

The ESEM approach to EFA allows the usual choices of rotations: orthogonal or oblique as well as bifactor.

The ESEM specification allows more general modeling including residual covariances and including other variables such as covariates. In fact, ESEM generalizes to SEM settings where there can be more than one EFA block and where EFA blocks can be combined with CFA blocks.



PSEM with GEOMIN Priors Input for H&S

```
ANALYSIS:        ESTIMATOR = MLR;
                 ! some of the below settings might
                 ! be needed:
                 ! ITERATIONS = 10000;
                 ! CONVERGENCE = 0.000001;
                 ! STARTS = 50;

MODEL:           ! label the 4*19=76 factor loadings
                 ! for which the GEOMIN rotation
                 ! should be applied:
                 f1-f4 BY visual-figurew*(a1-a76);
                 f1-f4@1;

MODEL PRIORS:    a1-a76 ~ GEOMIN(4,1.0); ! GEOMIN settings
                 ! are shown at the end of the Theory section
```

- The GEOMIN priors are applied to all factor loadings which gives the same rotated solution as the default for TYPE = EFA with 4 factors
- It is recommended to standardize the factor indicators before analysis so that the priors work optimally. This can be done using the STANDARDIZE = visual-figurew; option of the DEFINE command.

Slide 14 shows the input for PSEM with GEOMIN priors for the H&S data. This is our first PSEM example in that it uses priors. The commented lines explain the specifications.

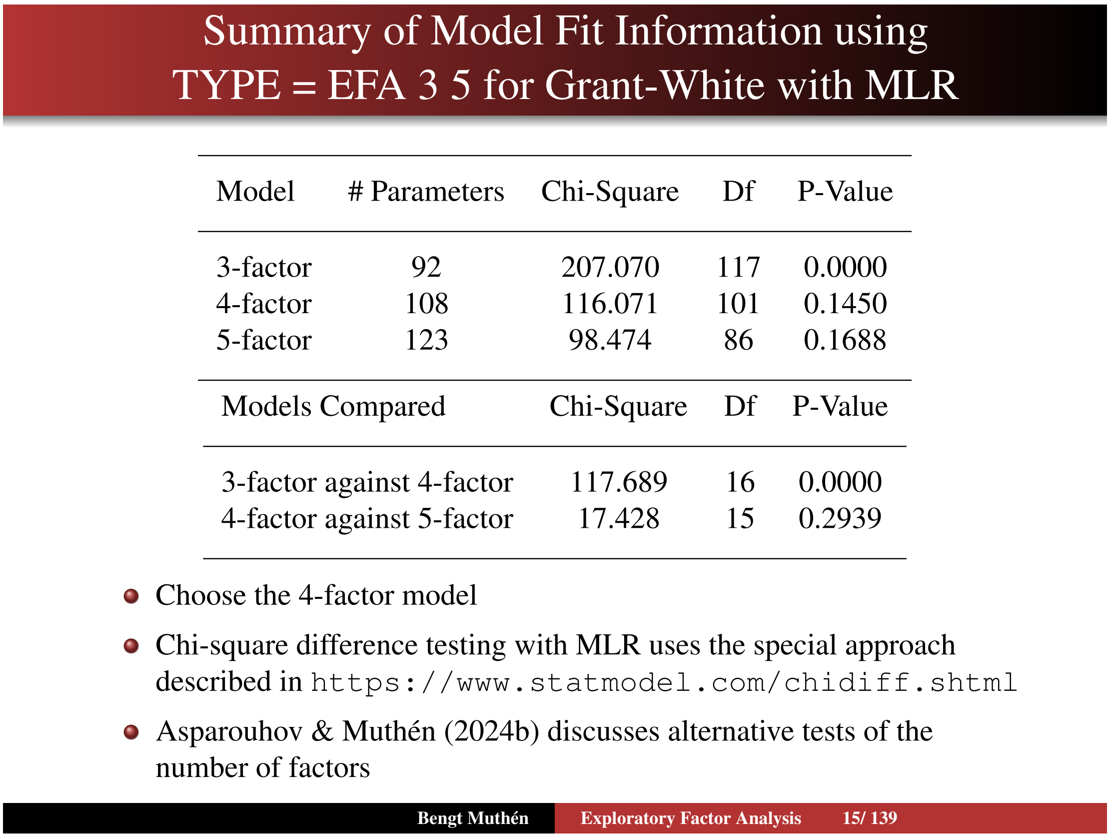
In the MODEL command, parameter labels are given to all 76 factor loadings.

The MODEL PRIORS command specifies GEOMIN priors for the loadings. This gives essentially the same rotated solution as the default GEOMIN rotation for TYPE = EFA with 4 factors.

The GEOMIN setting “4” refers to the number of factors and 1.0 refers to the prior variance or rather 1/the penalty weight. GEOMIN settings are given in the EFA Theory section.

The factor indicators in the Holzinger-Swineford data have very different variances. We recommend standardizing the factor indicators before analysis so that the priors work optimally. This can be done using the STANDARDIZE option in the DEFINE command. For EFA, the standardization does not affect the chi-square testing.

As we will see, the GEOMIN priors approach is very convenient in that it can be used for other applications such as second-order analysis, bi-factor analysis, EFA in structural equation modeling settings, and EFA in longitudinal settings.



Summary of Model Fit Information using TYPE = EFA 3 5 for Grant-White with MLR

Model	# Parameters	Chi-Square	Df	P-Value
3-factor	92	207.070	117	0.0000
4-factor	108	116.071	101	0.1450
5-factor	123	98.474	86	0.1688

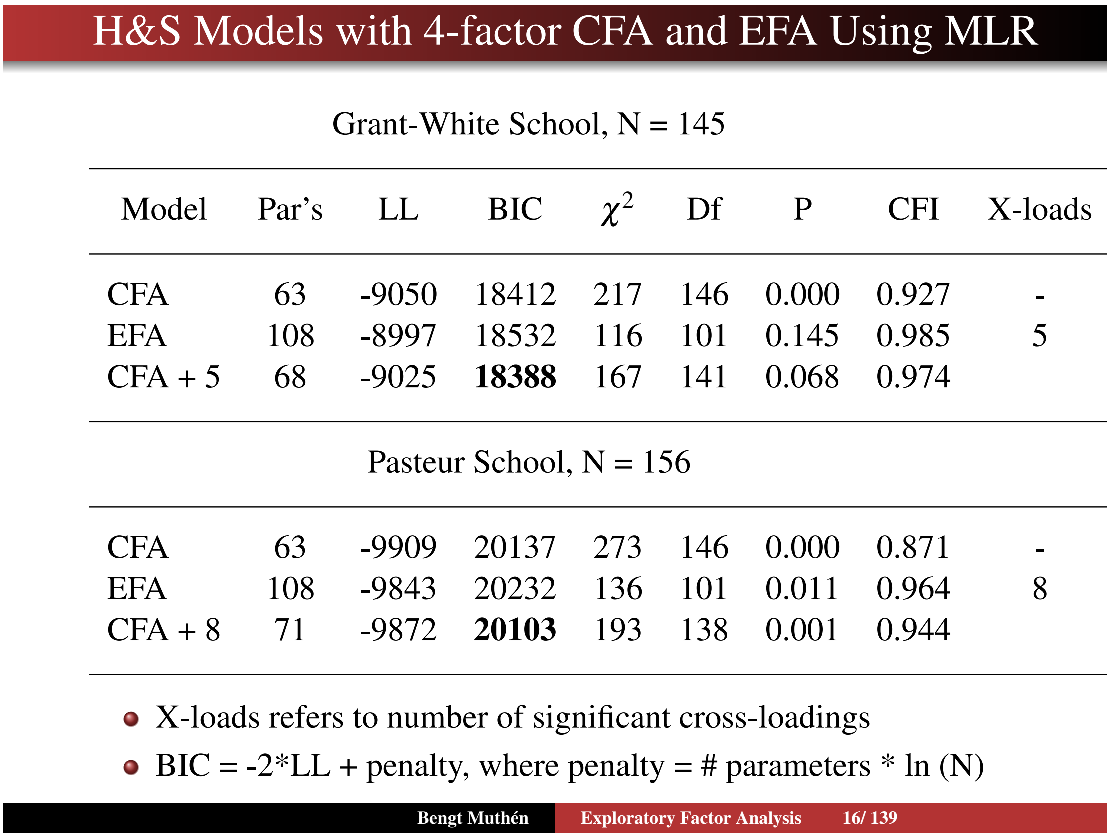
Models Compared	Chi-Square	Df	P-Value
3-factor against 4-factor	117.689	16	0.0000
4-factor against 5-factor	17.428	15	0.2939

- Choose the 4-factor model
- Chi-square difference testing with MLR uses the special approach described in `https://www.statmodel.com/chidiff.shtml`
- Asparouhov & Muthén (2024b) discusses alternative tests of the number of factors

Slide 15 shows the chi-square test of fit for 3, 4, and 5 factors. This is for the Grant-White school using MLR. Both a 4- and a 5-factor model fit well, but chi-square difference testing shows that 4 factors is sufficient in that it cannot be rejected when tested against 5 factors. So, we choose 4 factors.

With MLR, the chi-square difference testing uses the Satorra-Bentler approach that you find described on the Mplus website:
`https://www.statmodel.com/chidiff.shtml`

Alternative ways to test the number of factors in EFA is discussed in the Asparouhov & Muthén (2024) web note 25.



H&S Models with 4-factor CFA and EFA Using MLR

Grant-White School, N = 145

Model	Par's	LL	BIC	χ^2	Df	P	CFI	X-loads
CFA	63	-9050	18412	217	146	0.000	0.927	-
EFA	108	-8997	18532	116	101	0.145	0.985	5
CFA + 5	68	-9025	**18388**	167	141	0.068	0.974	

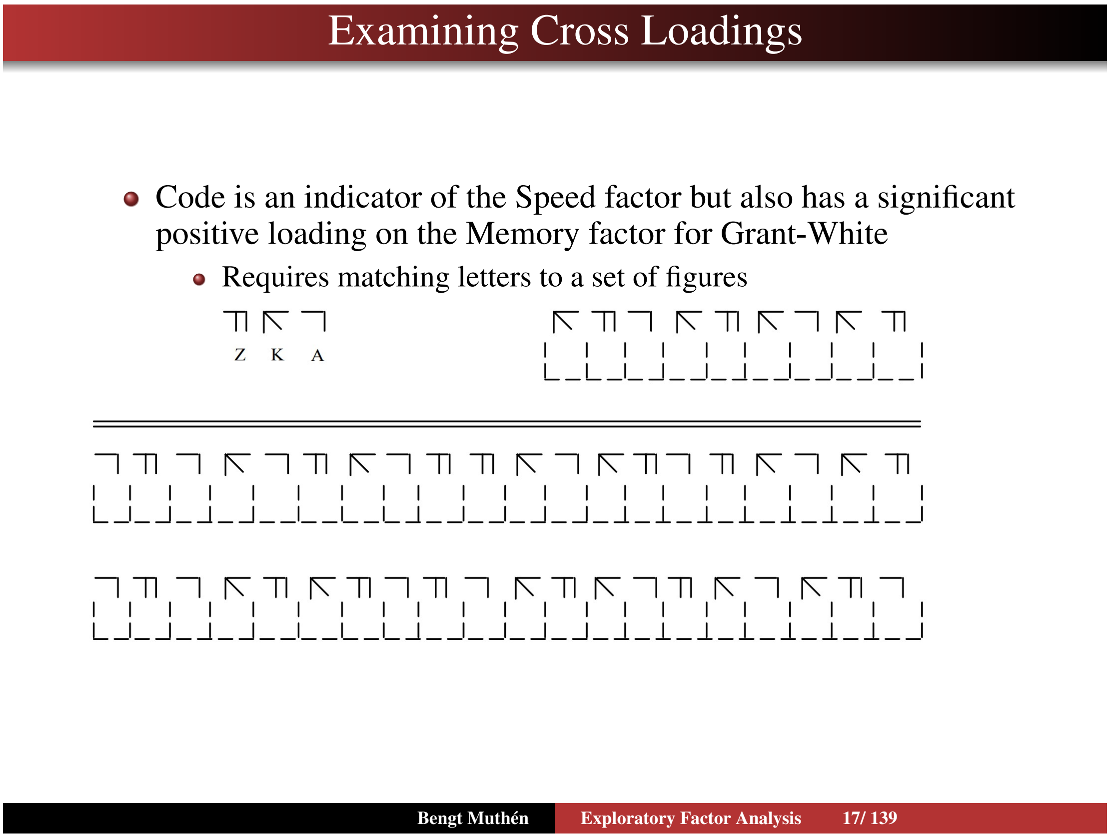
Pasteur School, N = 156

Model	Par's	LL	BIC	χ^2	Df	P	CFI	X-loads
CFA	63	-9909	20137	273	146	0.000	0.871	-
EFA	108	-9843	20232	136	101	0.011	0.964	8
CFA + 8	71	-9872	**20103**	193	138	0.001	0.944	

- X-loads refers to number of significant cross-loadings
- BIC = -2*LL + penalty, where penalty = # parameters * ln (N)

Slide 16 shows CFA and EFA test results for both schools. For both schools, EFA fits better in that its chi-square test gets a non-zero p-value unlike CFA. BIC, however, points to CFA. The BIC formula is shown at the bottom of the slide. This disagreement is not uncommon. In a sense, BIC shortchanges EFA. Not all EFA cross loadings are expected to be significant but these extra parameters are counted in the BIC penalty. As the table shows, EFA has a better log likelihood (LL) but many more parameters. The log likelihood improvement for EFA is not enough to compensate for the BIC penalty due to the large increase in number of parameters.

Out of the total of 57 cross loadings, Grant-White EFA has 5 that are significant and Pasteur has 8. Freeing these cross loadings in a modified CFA, the table shows that the best BIC is obtained for both schools. In this way, EFA provides an alternative to modification indices in terms of modifying a CFA. EFA may be closer to a well-fitting model than CFA so modifications of CFA based on EFA can be more likely to succeed than modifying the original CFA model by modification indices.



Examining Cross Loadings

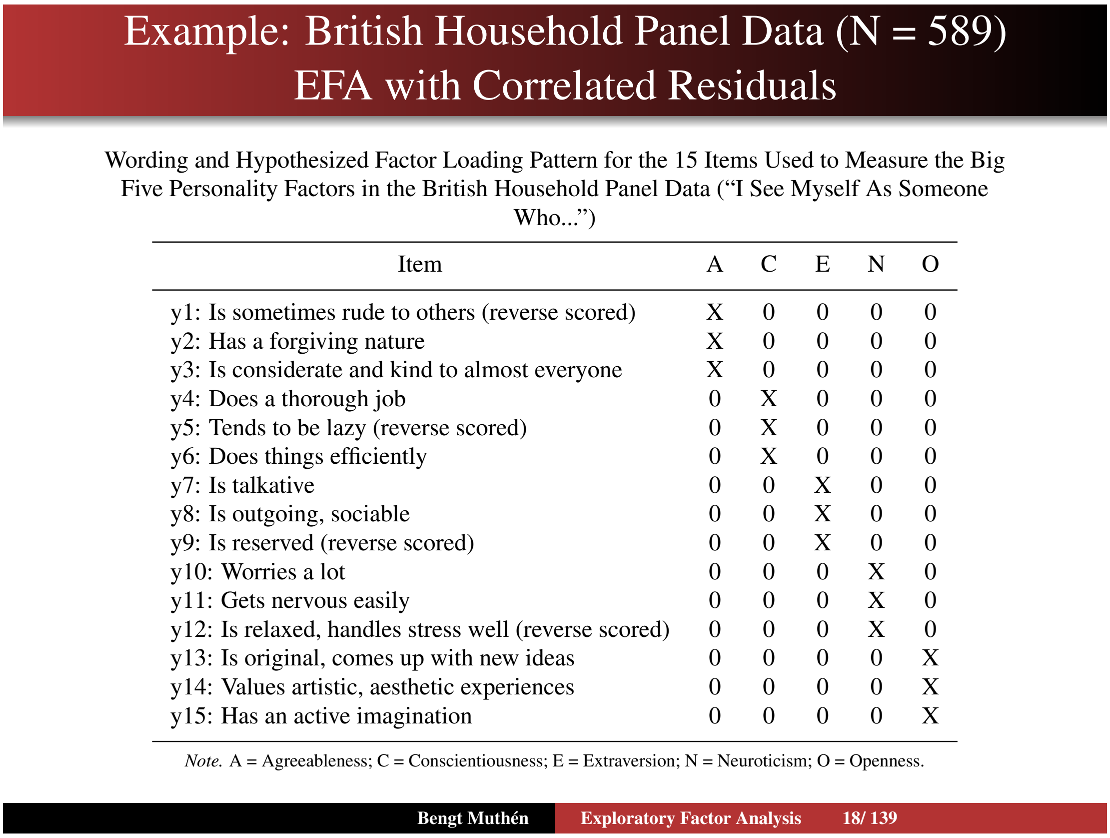
- Code is an indicator of the Speed factor but also has a significant positive loading on the Memory factor for Grant-White
  - Requires matching letters to a set of figures

Bengt Muthén Exploratory Factor Analysis 17/ 139

Slide 17 shows an example of a cross loading in the analysis of H&S. This is one of the 5 significant cross loadings from the EFA of Grant-White. The indicator called code measures the speed factor but also loads on the memory factor. This makes sense in that the icons for the Z, K, A letters have to be remembered when quickly filling in the letters under the icons.

The cross loading is added to the CFA by saying: memory BY code;

2.1.2 British Household Panel example. Correlated residuals



Example: British Household Panel Data (N = 589)
EFA with Correlated Residuals

Wording and Hypothesized Factor Loading Pattern for the 15 Items Used to Measure the Big Five Personality Factors in the British Household Panel Data ("I See Myself As Someone Who...")

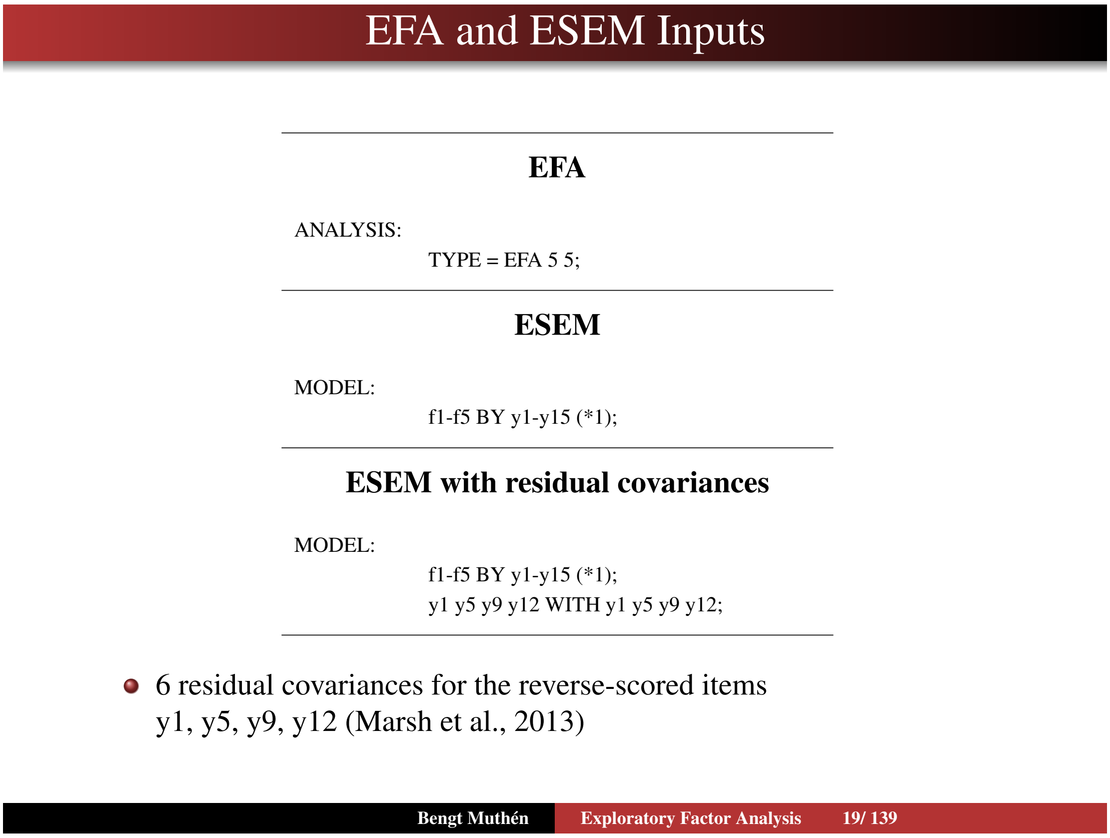
Item	A	C	E	N	O
y1: Is sometimes rude to others (reverse scored)	X	0	0	0	0
y2: Has a forgiving nature	X	0	0	0	0
y3: Is considerate and kind to almost everyone	X	0	0	0	0
y4: Does a thorough job	0	X	0	0	0
y5: Tends to be lazy (reverse scored)	0	X	0	0	0
y6: Does things efficiently	0	X	0	0	0
y7: Is talkative	0	0	X	0	0
y8: Is outgoing, sociable	0	0	X	0	0
y9: Is reserved (reverse scored)	0	0	X	0	0
y10: Worries a lot	0	0	0	X	0
y11: Gets nervous easily	0	0	0	X	0
y12: Is relaxed, handles stress well (reverse scored)	0	0	0	X	0
y13: Is original, comes up with new ideas	0	0	0	0	X
y14: Values artistic, aesthetic experiences	0	0	0	0	X
y15: Has an active imagination	0	0	0	0	X

Note. A = Agreeableness; C = Conscientiousness; E = Extraversion; N = Neuroticism; O = Openness.

Bengt Muthén Exploratory Factor Analysis 18/ 139

Slide 18 shows a different example using data from the British Household Panel. A simple loading pattern for 5 factors is hypothesized. The factor names are listed below the table. This measurement instrument has 4 indicators that have negative statements as opposed to all the other indicators so they are reverse scored.

This example illustrates EFA with correlated residuals using ESEM.



EFA and ESEM Inputs

EFA

```
ANALYSIS:
          TYPE = EFA 5 5;
```

ESEM

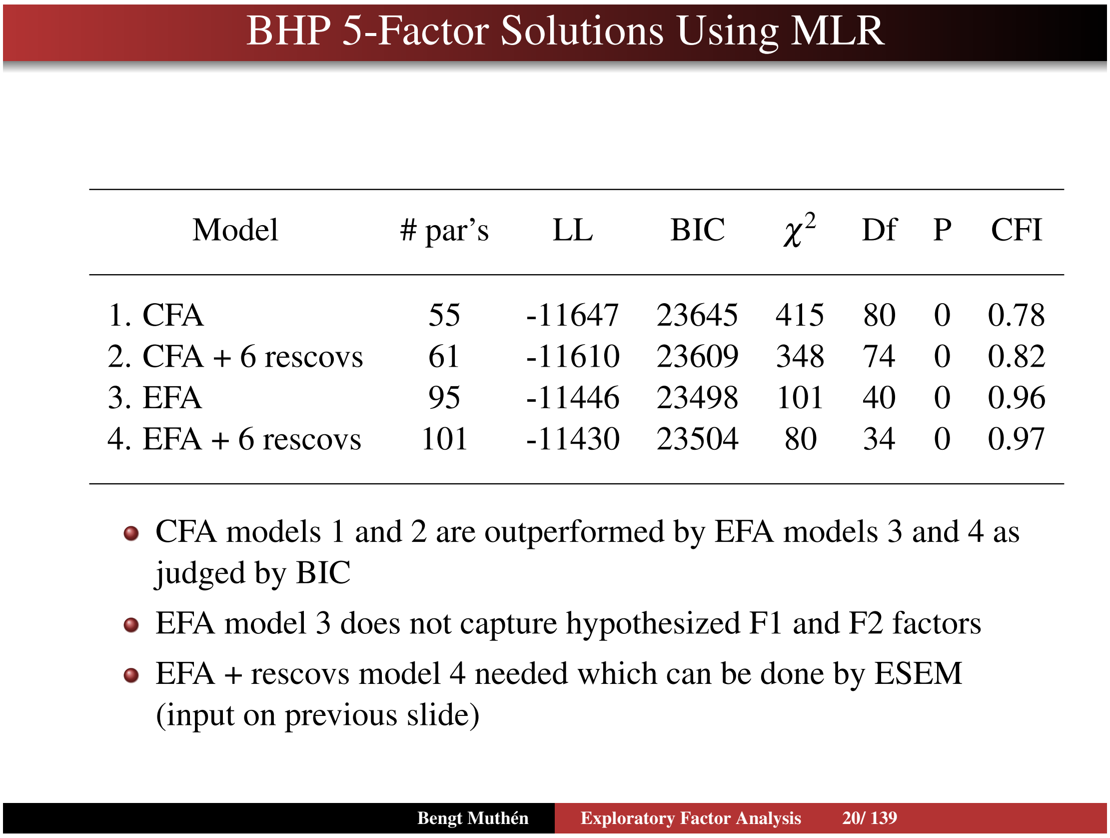
```
MODEL:
          f1-f5 BY y1-y15 (*1);
```

ESEM with residual covariances

```
MODEL:
          f1-f5 BY y1-y15 (*1);
          y1 y5 y9 y12 WITH y1 y5 y9 y12;
```

- 6 residual covariances for the reverse-scored items y1, y5, y9, y12 (Marsh et al., 2013)

Slide 19 shows EFA and ESEM inputs for this dataset. The bottom input shows the strength of ESEM in that it can include residual covariances in the model. These residual covariances are added for the reverse scored indicators per suggestion of the paper by Marsh et al. (2013) which analyzed this dataset. In this way, negatively worded indicators are allowed to correlate beyond what the factors would predict.



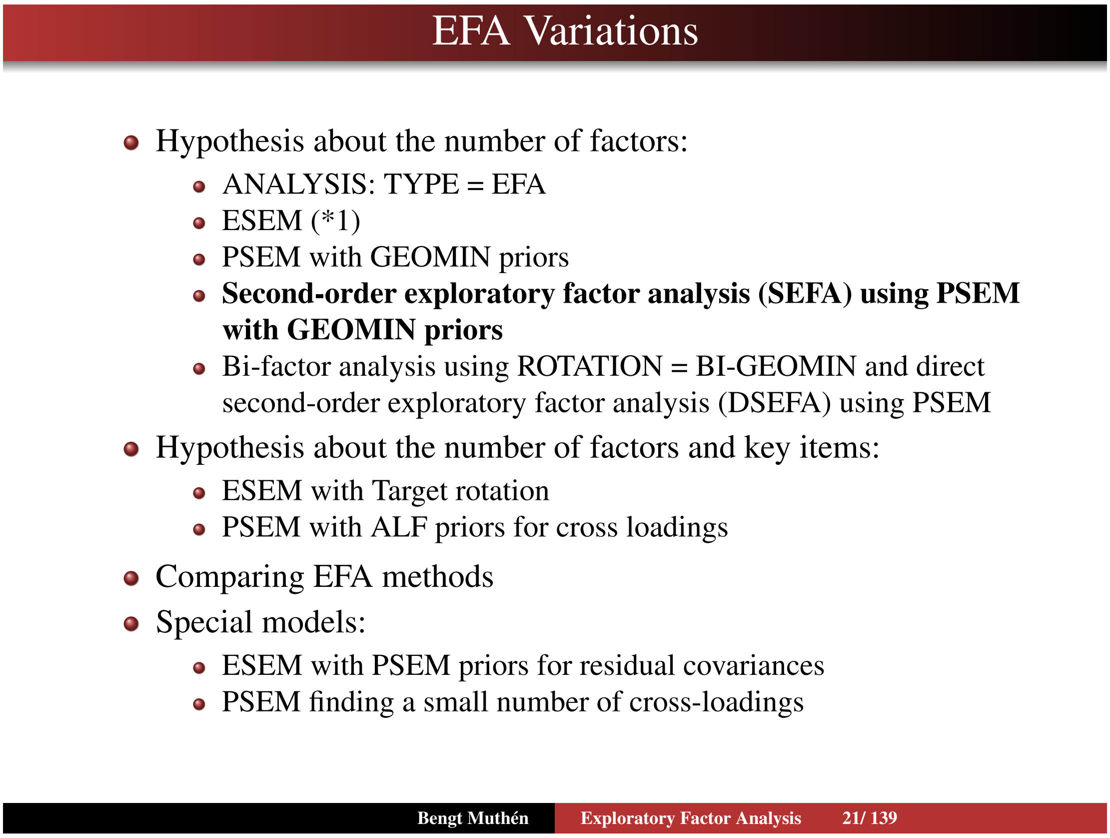
BHP 5-Factor Solutions Using MLR

Model	# par’s	LL	BIC	χ^2	Df	P	CFI
1. CFA	55	-11647	23645	415	80	0	0.78
2. CFA + 6 rescovs	61	-11610	23609	348	74	0	0.82
3. EFA	95	-11446	23498	101	40	0	0.96
4. EFA + 6 rescovs	101	-11430	23504	80	34	0	0.97

- CFA models 1 and 2 are outperformed by EFA models 3 and 4 as judged by BIC
- EFA model 3 does not capture hypothesized F1 and F2 factors
- EFA + rescovs model 4 needed which can be done by ESEM (input on previous slide)

Slide 20 shows the model fit for 4 different models. Comparing the CFA models 1 and 2, it is seen that the addition of the 6 residual covariances is warranted given better BIC and CFI. The CFA models are however outperformed by the two EFA models in terms of BIC. The EFA models also have better CFI. EFA model 3 has the best BIC but it doesn’t capture the first two factors. Not adding the residual covariances throws off the modeling. EFA model 4 which is done by ESEM gives a BIC value only slightly worse than for model 3 and is the model of choice.

2.1.3 Second-order EFA (SEFA) using PSEM with GEOMIN priors

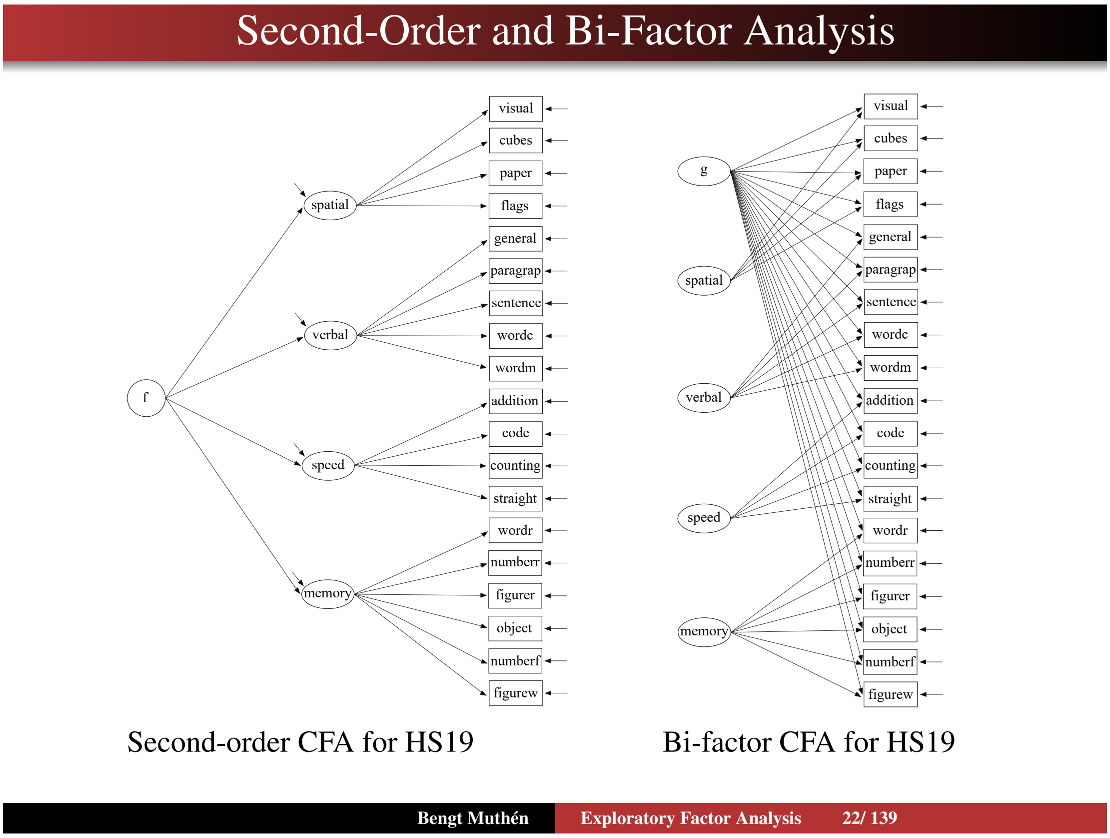


EFA Variations

- Hypothesis about the number of factors:
    - ANALYSIS: TYPE = EFA
    - ESEM (*1)
    - PSEM with GEOMIN priors
    - **Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors**
    - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- Hypothesis about the number of factors and key items:
    - ESEM with Target rotation
    - PSEM with ALF priors for cross loadings
- Comparing EFA methods
- Special models:
    - ESEM with PSEM priors for residual covariances
    - PSEM finding a small number of cross-loadings

Slide 21 returns to the overview of EFA Variations. We now turn to the topic of second-order factor analysis where we will also introduce the SEFA model using PSEM with GEOMIN priors.

The SEFA model also provides the basis for the new DSEFA approach to bi-factor modeling.



Second-Order and Bi-Factor Analysis

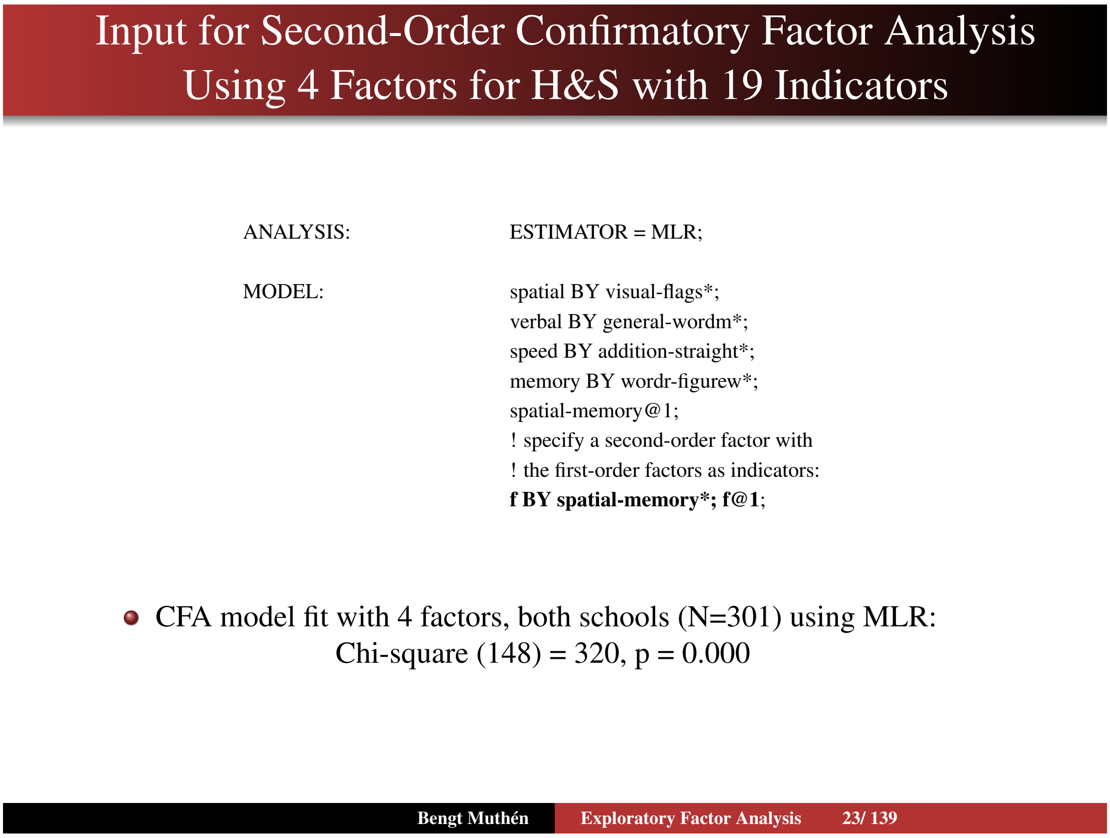
Second-order CFA for HS19 Bi-factor CFA for HS19

Bengt Muthén **Exploratory Factor Analysis** **22/ 139**

Slide 22 shows model diagrams for confirmatory factor analysis of a second-order factor model on the left and a bi-factor model on the right. The diagrams use the example of the H&S data discussed earlier with 19 tests used as factor indicators. Each of the models feature the 4 factors discussed earlier: spatial, verbal, speed, and memory. But the 4 factors play different roles in the two models. Both models consider if there is some sort of overall ability dimension that influences the performance on the tests. This could be a useful summary of ability either in its own right or in relations to background variables or prediction of future outcomes. The two models formulate this overall ability in different ways.

For the second-order model on the left, there is second-order factor f which influences the 4 first-order factors and therefore indirectly influences the indicators. The factor f is what all of the first-order factors have in common and it fully explains the correlations among the first-order factors.

For the bi-factor model on the right, there is a general factor that directly influences all the indicators. In addition, there are 4 specific factors that can be seen as residual factors, explaining correlations among the indicators beyond what the general factor can explain. All 5 factors in this model are typically specified to be uncorrelated.



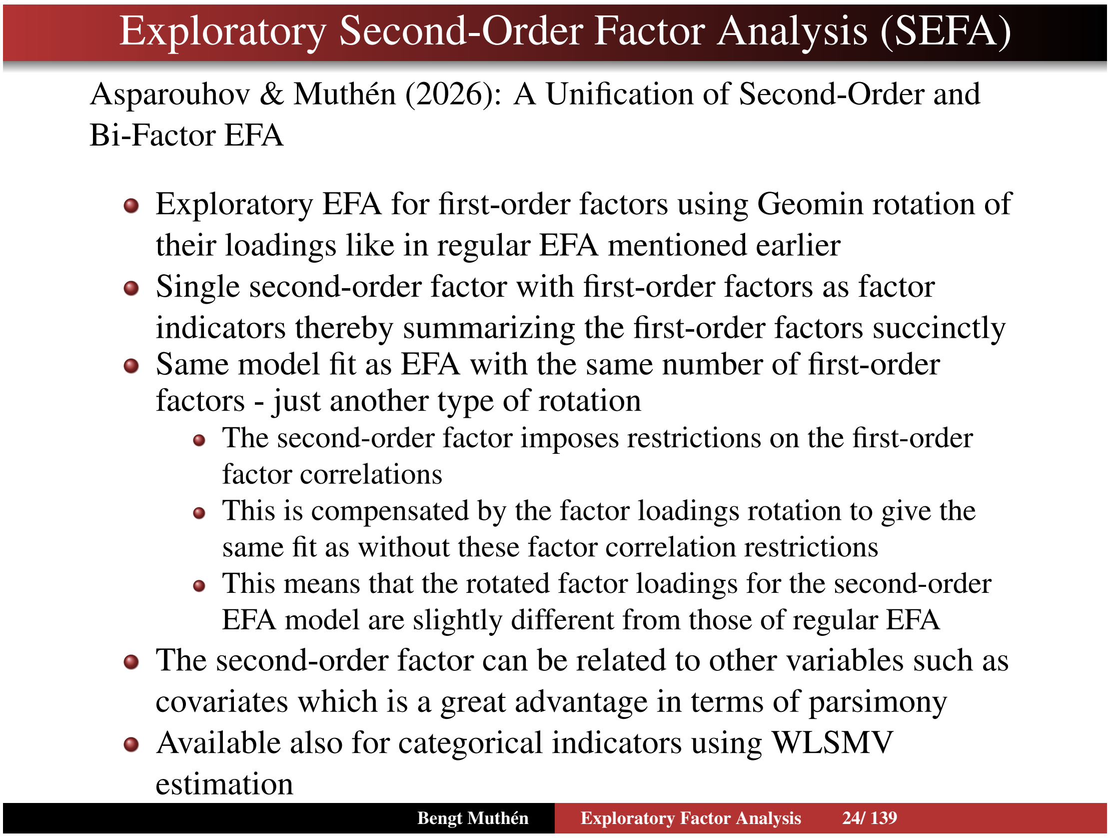
Input for Second-Order Confirmatory Factor Analysis Using 4 Factors for H&S with 19 Indicators

```
ANALYSIS:         ESTIMATOR = MLR;

MODEL:            spatial BY visual-flags*;
                  verbal BY general-wordm*;
                  speed BY addition-straight*;
                  memory BY wordr-figurew*;
                  spatial-memory@1;
                  ! specify a second-order factor with
                  ! the first-order factors as indicators:
                  f BY spatial-memory*; f@1;
```

- CFA model fit with 4 factors, both schools (N=301) using MLR:
  Chi-square (148) = 320, p = 0.000

Slide 23 shows the input for the second-order CFA model on the previous slide. All loadings are allowed to be free. The metric of the factors is set by fixing the factor variances at 1.

The CFA chi-square test of fit is shown at the bottom and indicates that the model fits poorly.



Exploratory Second-Order Factor Analysis (SEFA)

Asparouhov & Muthén (2026): A Unification of Second-Order and Bi-Factor EFA

- Exploratory EFA for first-order factors using Geomin rotation of their loadings like in regular EFA mentioned earlier
- Single second-order factor with first-order factors as factor indicators thereby summarizing the first-order factors succinctly
- Same model fit as EFA with the same number of first-order factors - just another type of rotation
  - The second-order factor imposes restrictions on the first-order factor correlations
  - This is compensated by the factor loadings rotation to give the same fit as without these factor correlation restrictions
  - This means that the rotated factor loadings for the second-order EFA model are slightly different from those of regular EFA
- The second-order factor can be related to other variables such as covariates which is a great advantage in terms of parsimony
- Available also for categorical indicators using WLSMV estimation

Slide 24 turns to exploratory second-order factor analysis, referred to as SEFA. This was introduced in our 2026 paper, A unification of second-order and bi-factor EFA.

SEFA is an example of PSEM analysis. The bullets show the features of SEFA.

Exploratory EFA for first-order factors uses Geomin rotation of their loadings like in regular EFA mentioned earlier.

There is a single second-order factor with first-order factors as factor indicators thereby summarizing the first-order factors succinctly.

SEFA has the same model fit as EFA with the same number of first-order factors; SEFA is just another type of rotation where:

- The second-order factor imposes restrictions on the first-order factor correlations.

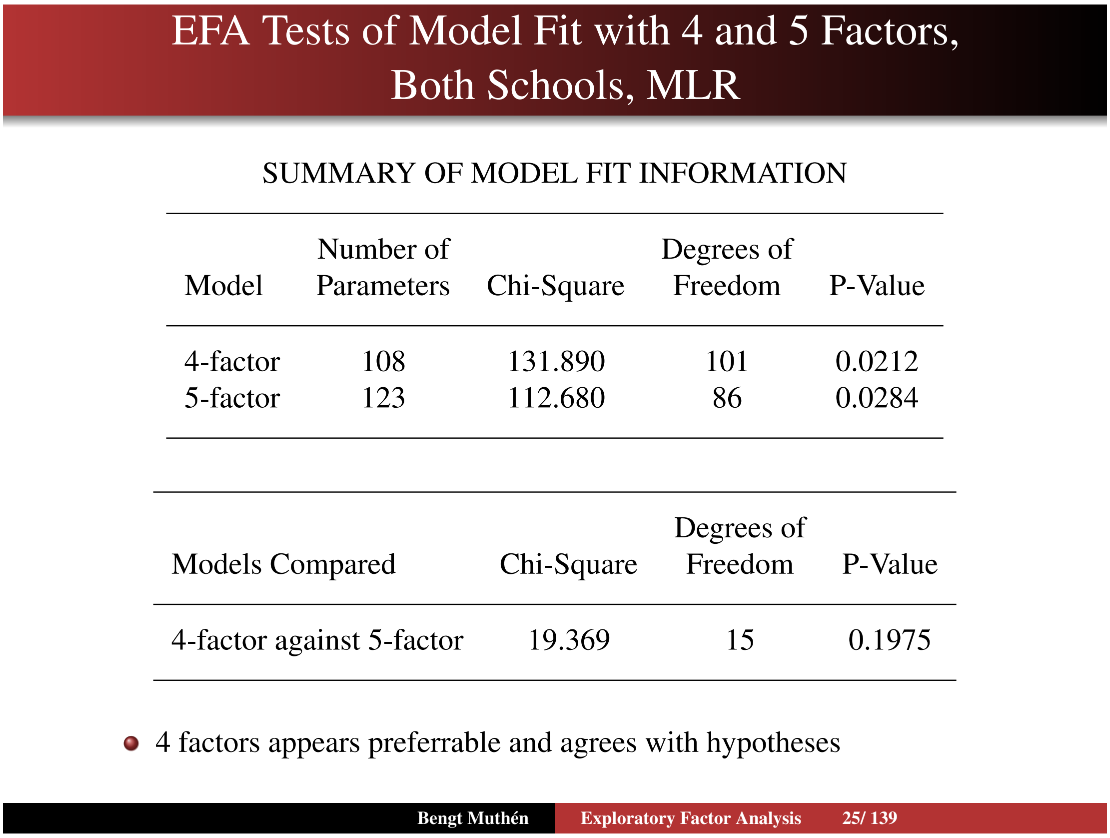
- This is compensated by the factor loadings rotation to give the same fit as without these factor correlation restrictions.

- This means that the rotated factor loadings for the second-order EFA model are slightly different from those of regular EFA.

The second-order factor can be related to other variables such as covariates which is a great advantage in terms of parsimony.

Pre-standardization of indicators important with widely varying variances so that the priors are used optimally.

SEFA is available not only for maximum-likelihood estimation of continuous indicators but also for categorical indicators using WLSMV estimation.



EFA Tests of Model Fit with 4 and 5 Factors, Both Schools, MLR

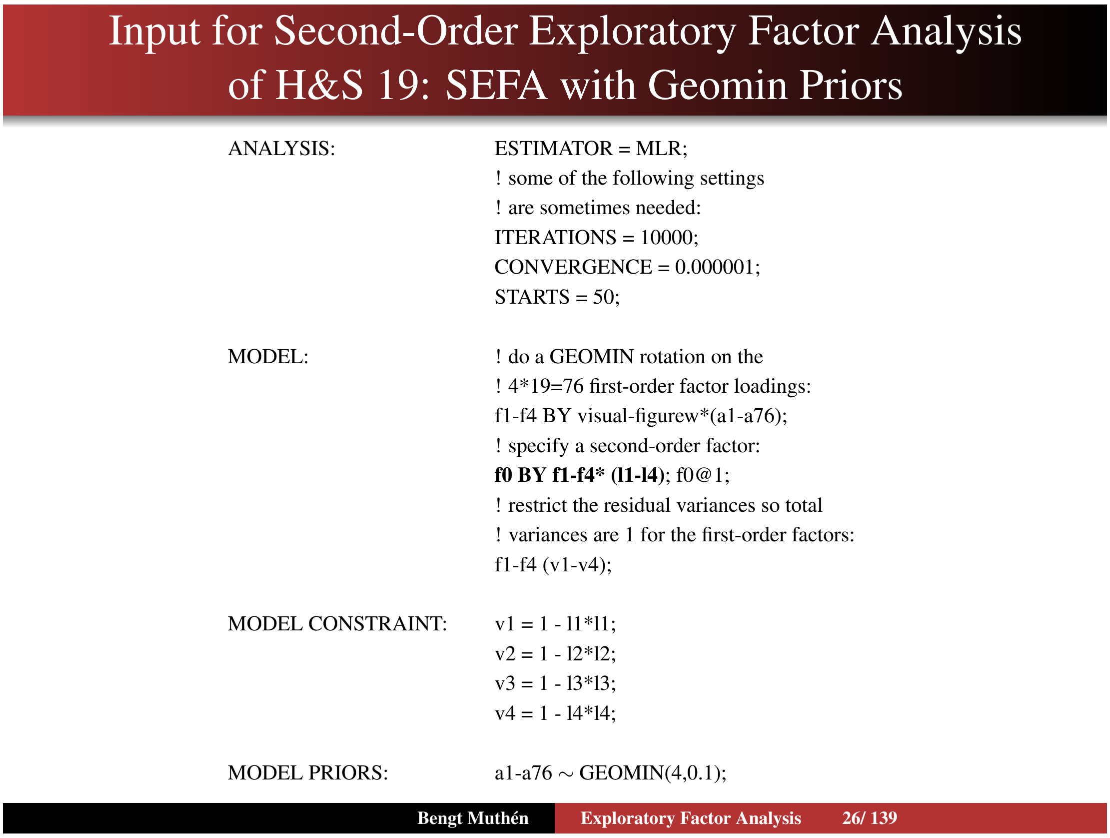
SUMMARY OF MODEL FIT INFORMATION

Model	Number of Parameters	Chi-Square	Degrees of Freedom	P-Value
4-factor	108	131.890	101	0.0212
5-factor	123	112.680	86	0.0284

Models Compared	Chi-Square	Degrees of Freedom	P-Value
4-factor against 5-factor	19.369	15	0.1975

- 4 factors appears preferrable and agrees with hypotheses

Slide 25 shows tests of model fit for EFA. We have seen this type of EFA output before but now it is presented for both schools (N = 301). These EFA results are shown because as stated on the previous slide, EFA has the same model fit as the SEFA exploratory second-order factor analysis model with the same number of first-order factors. SEFA simply results in a different rotation, a rotation that assumes a second-order factor. It is seen that 4 factors is preferable also for the combined sample.

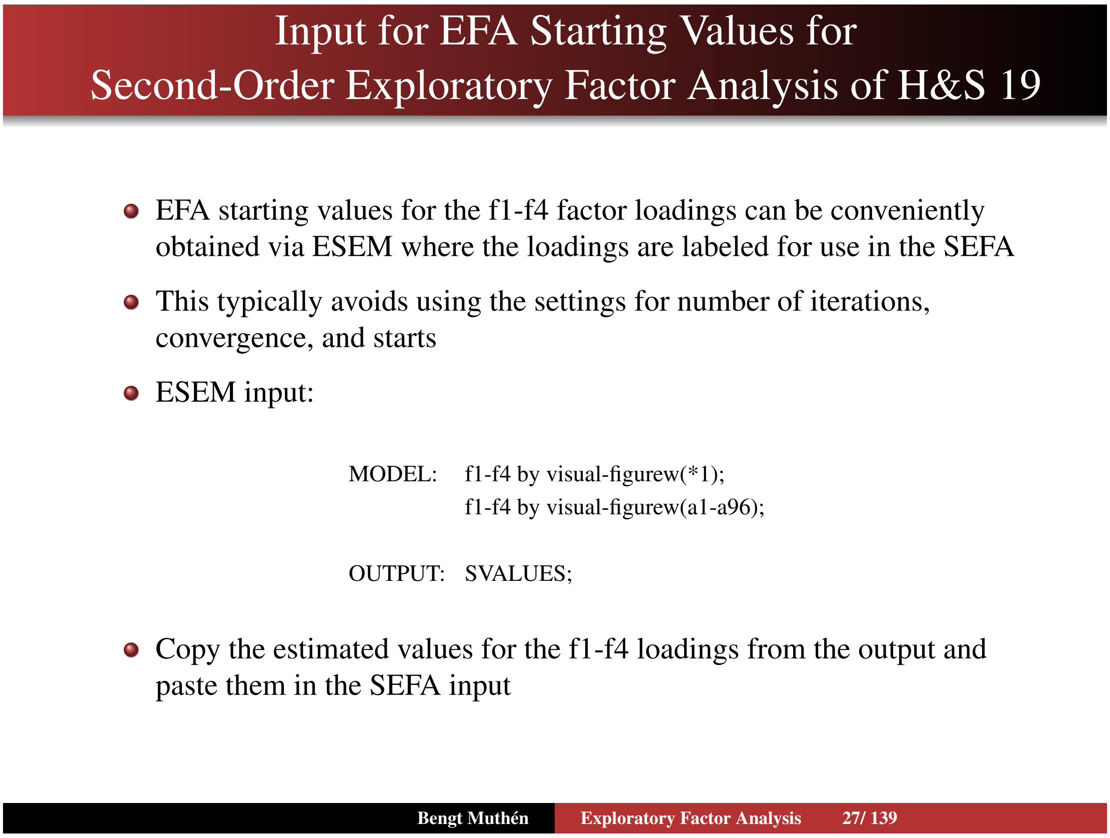


Input for Second-Order Exploratory Factor Analysis of H&S 19: SEFA with Geomin Priors

```
ANALYSIS:              ESTIMATOR = MLR;
                       ! some of the following settings
                       ! are sometimes needed:
                       ITERATIONS = 10000;
                       CONVERGENCE = 0.000001;
                       STARTS = 50;

MODEL:                 ! do a GEOMIN rotation on the
                       ! 4*19=76 first-order factor loadings:
                       f1-f4 BY visual-figurew*(a1-a76);
                       ! specify a second-order factor:
                       f0 BY f1-f4* (l1-l4); f0@1;
                       ! restrict the residual variances so total
                       ! variances are 1 for the first-order factors:
                       f1-f4 (v1-v4);

MODEL CONSTRAINT:      v1 = 1 - l1*l1;
                       v2 = 1 - l2*l2;
                       v3 = 1 - l3*l3;
                       v4 = 1 - l4*l4;

MODEL PRIORS:          a1-a76 ~ GEOMIN(4,0.1);
```

Slide 26 shows the input for SEFA as applied to the H&S dataset with 4 first-order factors. This is what we call the manual specification of the model. In Mplus Version 9.1 the input is to a large extent simplified by automatically doing most of the setup. This simplified input will also be shown.

The commented lines explain the input for the manual specification. Importantly, the GEOMIN rotation is applied to the first-order factors as seen in the labeling of the loadings and also in the MODEL PRIORS command. As pointed out in the paper, the residual variances need to be restricted so the total variance of each first-order factors is 1. This is carried out by the MODEL CONSTRAINT command.



Input for EFA Starting Values for Second-Order Exploratory Factor Analysis of H&S 19

- EFA starting values for the f1-f4 factor loadings can be conveniently obtained via ESEM where the loadings are labeled for use in the SEFA
- This typically avoids using the settings for number of iterations, convergence, and starts
- ESEM input:

```
MODEL:   f1-f4 by visual-figurew(*1);
         f1-f4 by visual-figurew(a1-a96);

OUTPUT:  SVALUES;
```

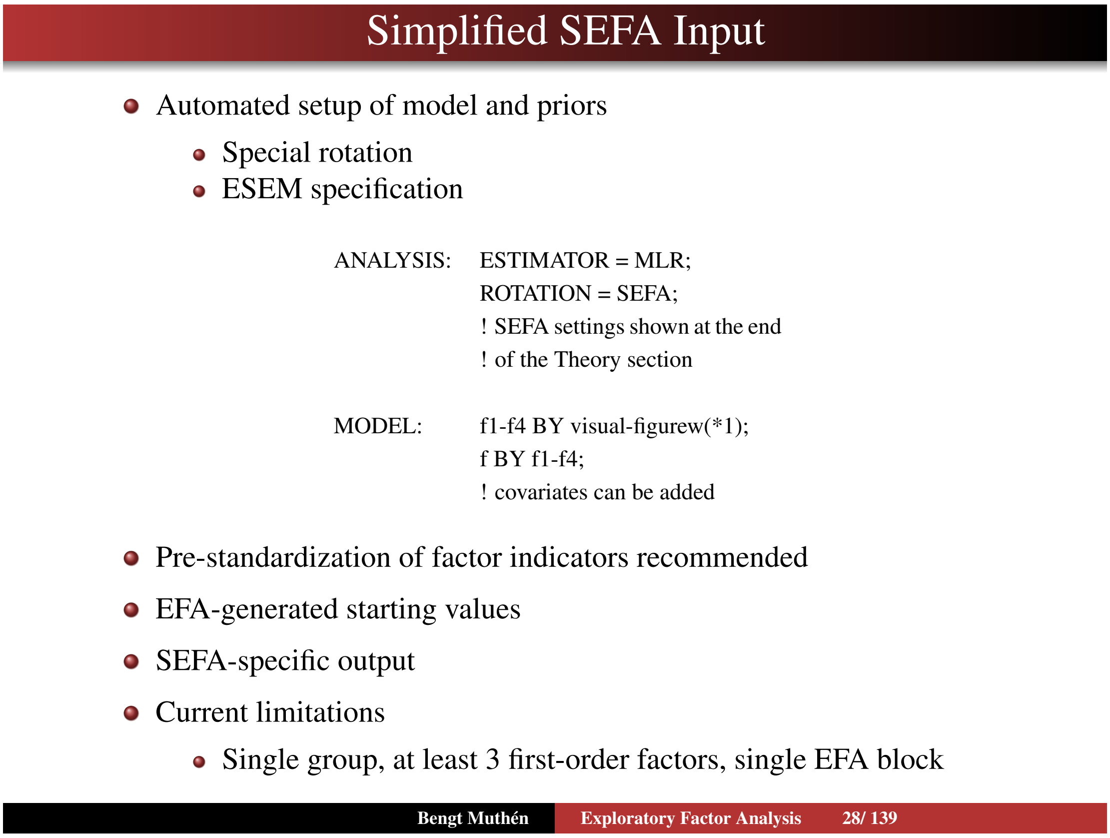
- Copy the estimated values for the f1-f4 loadings from the output and paste them in the SEFA input

Slide 27 shows the input for getting SEFA starting values using EFA in ESEM form. This is not needed for the simplified input but is shown here for completeness.

Starting values are helpful for avoiding the ANALYSIS specifications on the previous slide for iterations, convergence, and starts.

The ESEM input gives labels for the loadings which will be printed in the output by requesting SVALUES.

The SVALUES output for the loadings is then used in the SEFA input.



Simplified SEFA Input

- Automated setup of model and priors
  - Special rotation
  - ESEM specification

```
ANALYSIS:   ESTIMATOR = MLR;
            ROTATION = SEFA;
            ! SEFA settings shown at the end
            ! of the Theory section

MODEL:      f1-f4 BY visual-figurew(*1);
            f BY f1-f4;
            ! covariates can be added
```

- Pre-standardization of factor indicators recommended
- EFA-generated starting values
- SEFA-specific output
- Current limitations
  - Single group, at least 3 first-order factors, single EFA block

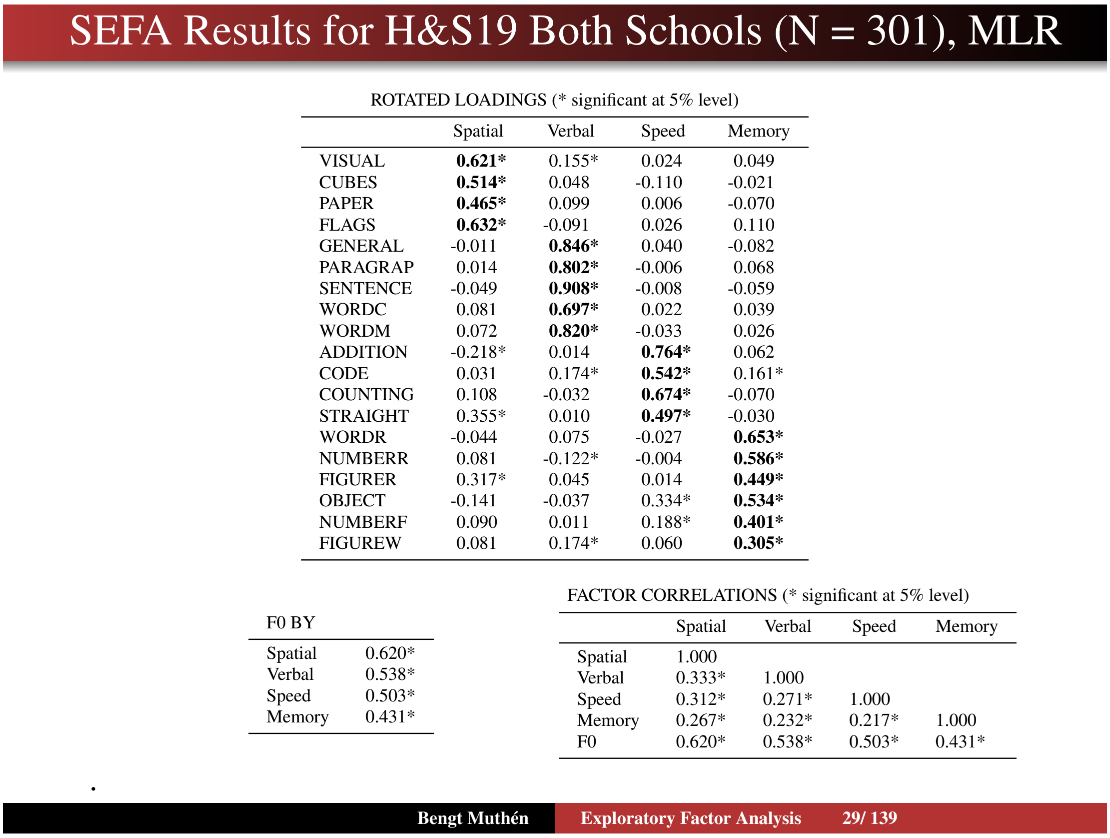
Slide 28 shows the simplified, automated input for SEFA replacing the previous two slides. The automated approach is activated by the ROTATION = SEFA statement, emphasizing that SEFA implies a specific rotation. Special settings can be given in parentheses, such as SEFA (0.1) corresponding to the GEOMIN setting of slide 26. The settings are described in the EFA Theory section.

The first-order factors are defined using ESEM style input. The added second-order factor is specified in usual style. Covariates can be added.

Standardization of the indicators is recommended when the indicators have widely varying variances.

The analysis uses starting values from an EFA analysis with the same number of factors as the number of first-order factors in the SEFA.

The automated approach to SEFA is currently limited to a single group with at least 3 first-order factors, as well as a single EFA block for ESEM.



SEFA Results for H&S19 Both Schools (N = 301), MLR

ROTATED LOADINGS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
VISUAL	**0.621***	0.155*	0.024	0.049
CUBES	**0.514***	0.048	-0.110	-0.021
PAPER	**0.465***	0.099	0.006	-0.070
FLAGS	**0.632***	-0.091	0.026	0.110
GENERAL	-0.011	**0.846***	0.040	-0.082
PARAGRAP	0.014	**0.802***	-0.006	0.068
SENTENCE	-0.049	**0.908***	-0.008	-0.059
WORDC	0.081	**0.697***	0.022	0.039
WORDM	0.072	**0.820***	-0.033	0.026
ADDITION	-0.218*	0.014	**0.764***	0.062
CODE	0.031	0.174*	**0.542***	0.161*
COUNTING	0.108	-0.032	**0.674***	-0.070
STRAIGHT	0.355*	0.010	**0.497***	-0.030
WORDR	-0.044	0.075	-0.027	**0.653***
NUMBERR	0.081	-0.122*	-0.004	**0.586***
FIGURER	0.317*	0.045	0.014	**0.449***
OBJECT	-0.141	-0.037	0.334*	**0.534***
NUMBERF	0.090	0.011	0.188*	**0.401***
FIGUREW	0.081	0.174*	0.060	**0.305***

F0 BY	
Spatial	0.620*
Verbal	0.538*
Speed	0.503*
Memory	0.431*

FACTOR CORRELATIONS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
Spatial	1.000			
Verbal	0.333*	1.000		
Speed	0.312*	0.271*	1.000	
Memory	0.267*	0.232*	0.217*	1.000
F0	0.620*	0.538*	0.503*	0.431*

.

Slide 29 shows the estimated factor loadings for the 4 first-order factors Spatial, Verbal, Speed, and Memory. All factors have large, significant loadings in the pattern that was hypothesized.

Bottom left shows that the first-order factors all have large, significant loadings on the second-order factor with largest loading for Spatial and smallest for Memory. The correlations among all 5 factors is also shown. We see that the first-order factors have moderate correlations. Of course, if they have small correlations, it doesn't make much sense to do SEFA.

As a footnote, the Spatial - Memory labeling of the columns of the factor loading matrix on slide 29 is not provided by Mplus but done afterwards based on interpretation. The Mplus output gives the f1-f4 labeling used in the input.

Comparison with Regular EFA 4-Factor Oblique Solution

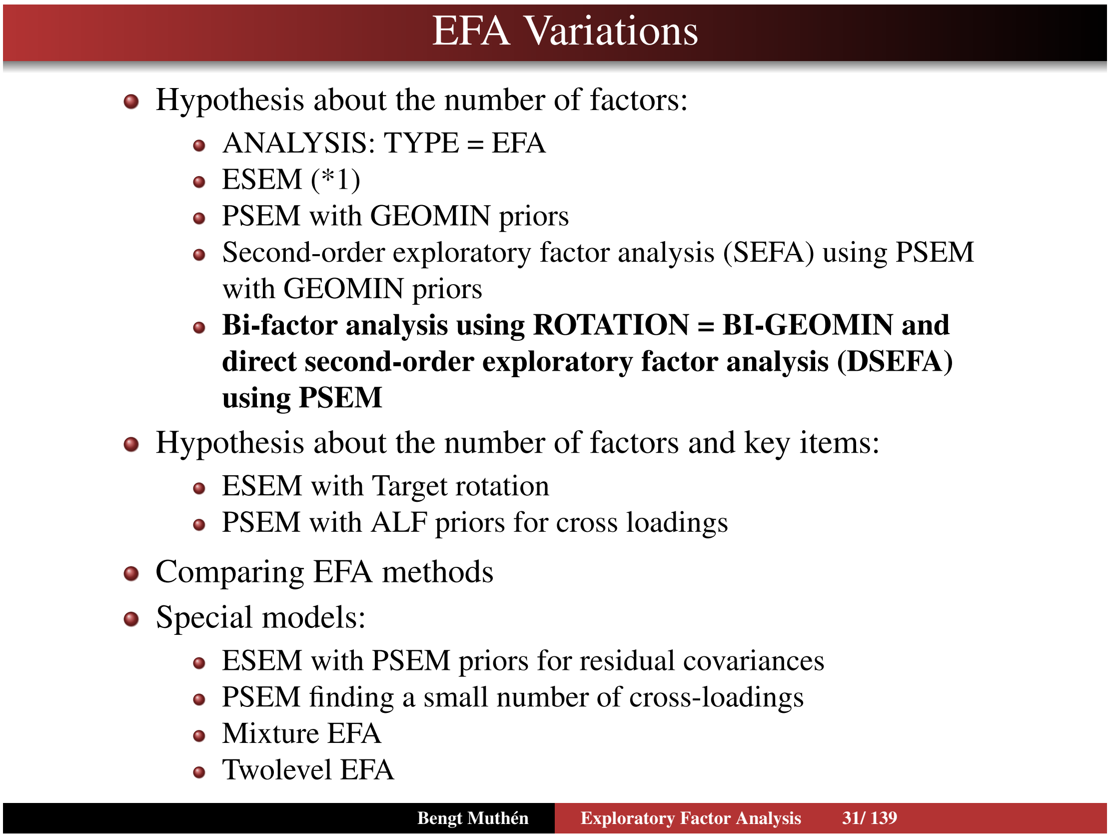
- Same model fit as regular 4-factor EFA
- For this example, the regular 4-factor EFA with the default oblique roations has similar factor loadings and factor correlations
  - Same number of significant cross-loadings and cross-loadings larger than 0.2 in absolute value
- The second-order factor can be related to other variables such as covariates

Slide 30 reminds us that the SEFA model has the same fit as a regular 4-factor EFA.

The default oblique rotation of EFA has similar factors loadings and factor correlations to those of SEFA. It has the same number of significant cross loadings and cross loadings larger than 0.2. This means that nothing is lost in terms of simplicity relative to regular EFA, but what is gained is the second-order factor summary of abilities.

The second-order factor can then be related to background variables and be used for predicting other outcomes.

2.1.4 Bi-factor analysis using ROTATION=BI-GEOMIN and direct effects second-order EFA (DSEFA) using PSEM

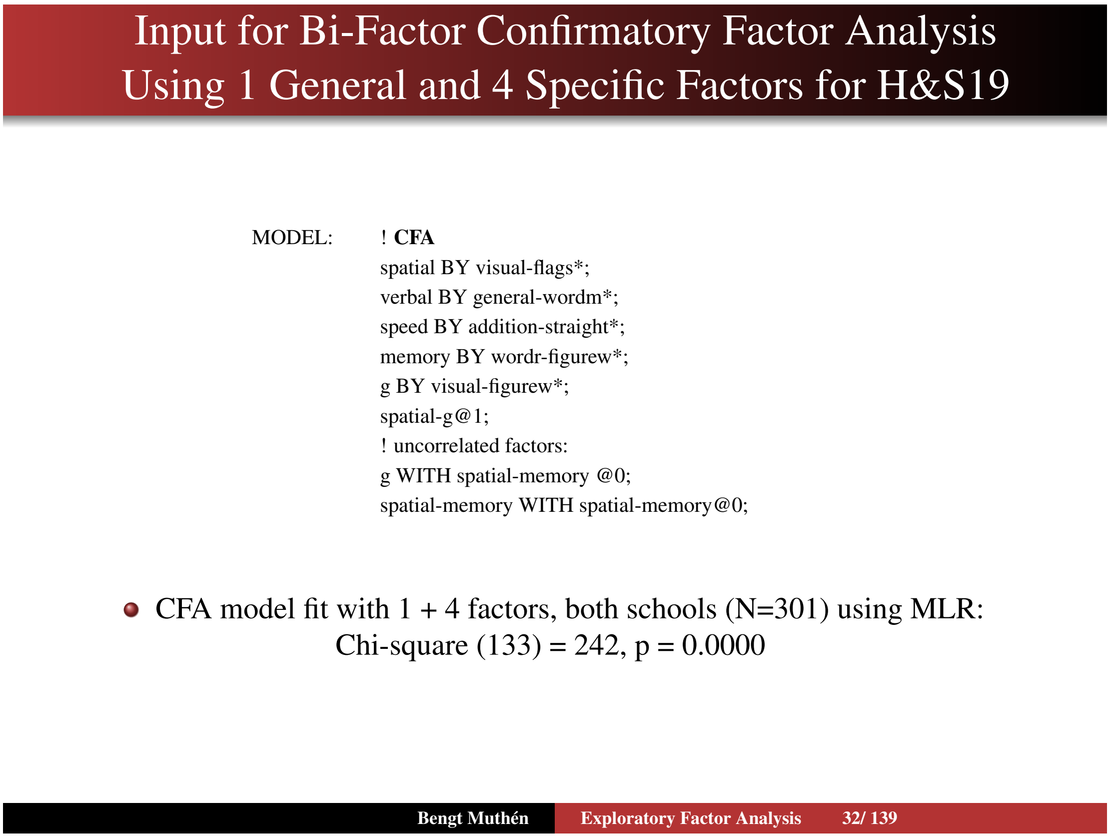


EFA Variations

- Hypothesis about the number of factors:
  - ANALYSIS: TYPE = EFA
  - ESEM (*1)
  - PSEM with GEOMIN priors
  - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
  - **Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM**
- Hypothesis about the number of factors and key items:
  - ESEM with Target rotation
  - PSEM with ALF priors for cross loadings
- Comparing EFA methods
- Special models:
  - ESEM with PSEM priors for residual covariances
  - PSEM finding a small number of cross-loadings
  - Mixture EFA
  - Twolevel EFA

Slide 31 once again shows the outline of the various EFA Variations. We now turn to bi-factor analysis. Both old and new, improved techniques are discussed.

The focus is on the new method of direct second-order exploratory factor analysis (DSEFA), which uses PSEM.

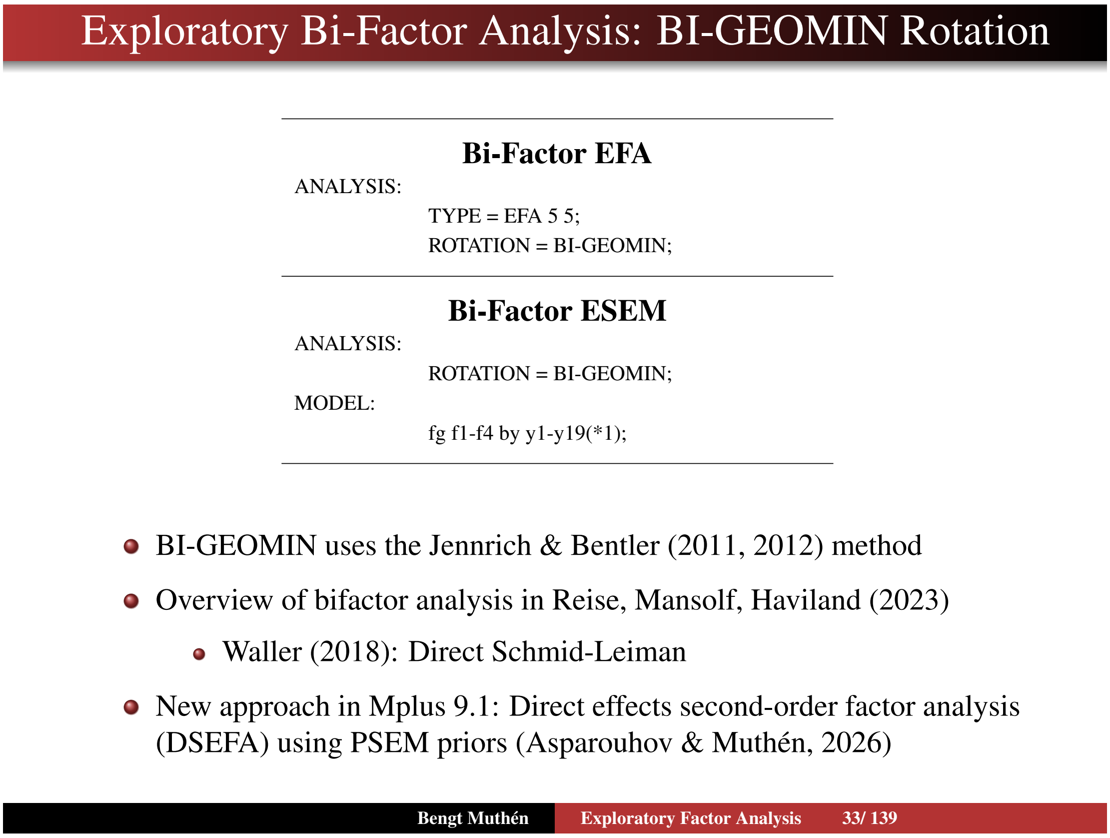


Input for Bi-Factor Confirmatory Factor Analysis Using 1 General and 4 Specific Factors for H&S19

```
MODEL:      ! CFA
            spatial BY visual-flags*;
            verbal BY general-wordm*;
            speed BY addition-straight*;
            memory BY wordr-figurew*;
            g BY visual-figurew*;
            spatial-g@1;
            ! uncorrelated factors:
            g WITH spatial-memory @0;
            spatial-memory WITH spatial-memory@0;
```

- CFA model fit with 1 + 4 factors, both schools (N=301) using MLR:
  Chi-square (133) = 242, p = 0.0000

Slide 32 shows the input for the bi-factor CFA model for the H&S19 data that we saw the model diagram for earlier. The 5 factors are specified to be uncorrelated in line with the diagram. The model fit is shown at the bottom.



Exploratory Bi-Factor Analysis: BI-GEOMIN Rotation

Bi-Factor EFA

```
ANALYSIS:
                TYPE = EFA 5 5;
                ROTATION = BI-GEOMIN;
```

Bi-Factor ESEM

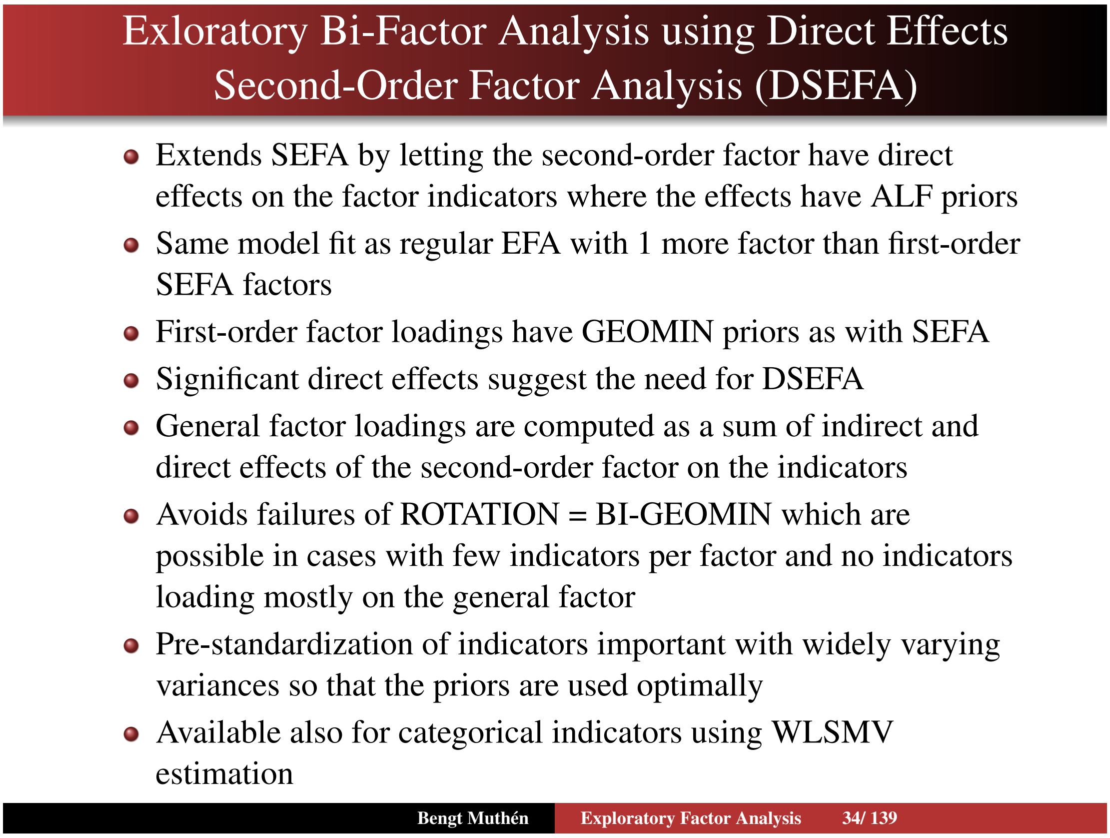
```
ANALYSIS:
                ROTATION = BI-GEOMIN;
MODEL:
                fg f1-f4 by y1-y19(*1);
```

- BI-GEOMIN uses the Jennrich & Bentler (2011, 2012) method
- Overview of bifactor analysis in Reise, Mansolf, Haviland (2023)
  - Waller (2018): Direct Schmid-Leiman
- New approach in Mplus 9.1: Direct effects second-order factor analysis (DSEFA) using PSEM priors (Asparouhov & Muthén, 2026)

Slide 33 shows input for exploratory bi-factor analysis using TYPE = EFA as well as using the equivalent ESEM specification. For ESEM, the general factor is called fg. Alternatively, this could be written as f1-f5 BY y1-y19(*1).

Both approaches use the BI-GEOMIN rotation, which is based on the Jennrich-Bentler Gradient Projection Algorithm (the Psychometrika references are given in the Reference section).

A very helpful overview of bi-factor analysis, both confirmatory and exploratory, is given in the chapter by Reise and others in the SEM handbook of 2003 (see reference section). That chapter also refers to a paper by Waller which inspired the new DSEFA approach in Mplus discussed in Asparouhov- Muthén (2026) and presented next. The DSEFA approach uses PSEM priors on direct effects from a second-order factor to the factor indicators to arrive at a new type of bi-factor analysis.



Exloratory Bi-Factor Analysis using Direct Effects Second-Order Factor Analysis (DSEFA)

- Extends SEFA by letting the second-order factor have direct effects on the factor indicators where the effects have ALF priors
- Same model fit as regular EFA with 1 more factor than first-order SEFA factors
- First-order factor loadings have GEOMIN priors as with SEFA
- Significant direct effects suggest the need for DSEFA
- General factor loadings are computed as a sum of indirect and direct effects of the second-order factor on the indicators
- Avoids failures of ROTATION = BI-GEOMIN which are possible in cases with few indicators per factor and no indicators loading mostly on the general factor
- Pre-standardization of indicators important with widely varying variances so that the priors are used optimally
- Available also for categorical indicators using WLSMV estimation

Slide 34 shows the features of DSEFA.

It extends SEFA by letting the second-order factor have direct effects on the factor indicators where the effects have ALF priors.

It has the same model fit as regular EFA with 1 more factor than the number of first-order SEFA factors.

First-order factor loadings have GEOMIN priors as with SEFA.

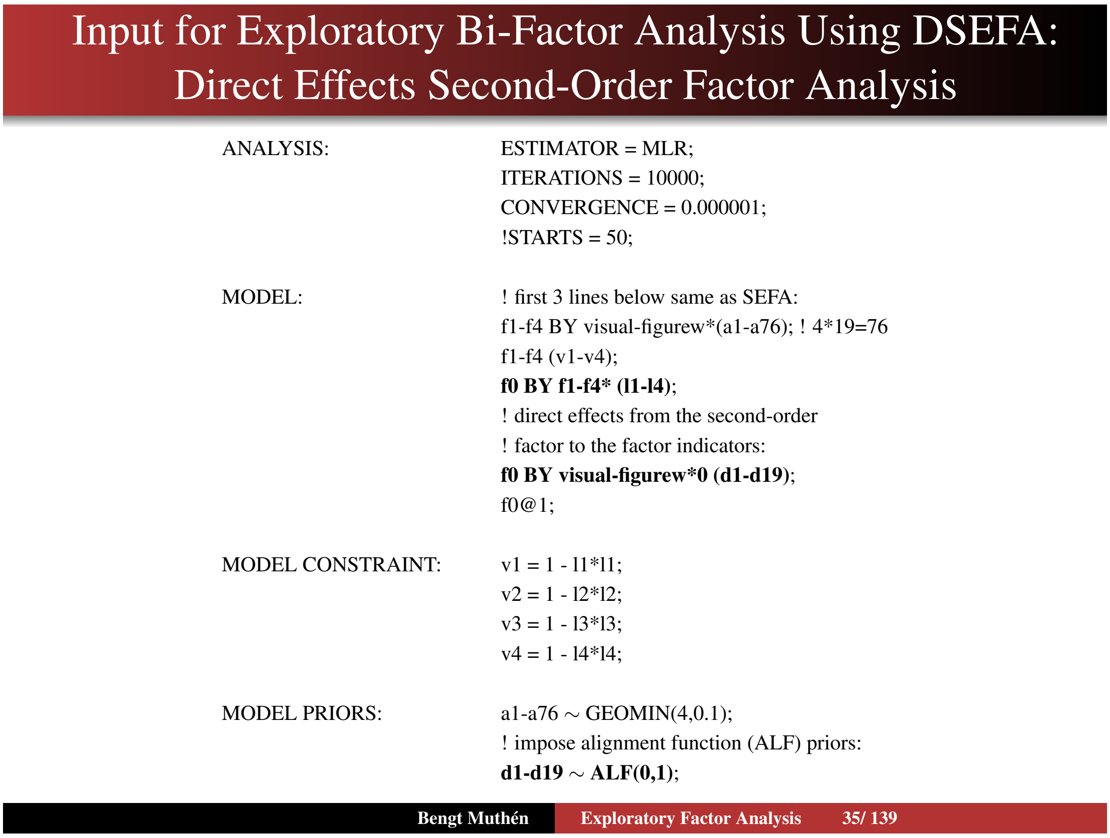
Significant direct effects suggest the need for DSEFA.

General factor loadings are computed as a sum of indirect and direct effects of the second-order factor on the indicators.

DSEFA avoids failures of ROTATION = BI-GEOMIN which are possible in cases with few indicators per factor and no indicators loading mostly on the general factor.

Pre-standardization is important with widely varying indicator variances.

DSEFA is also available for categorical indicators using the WLSMV estimator.



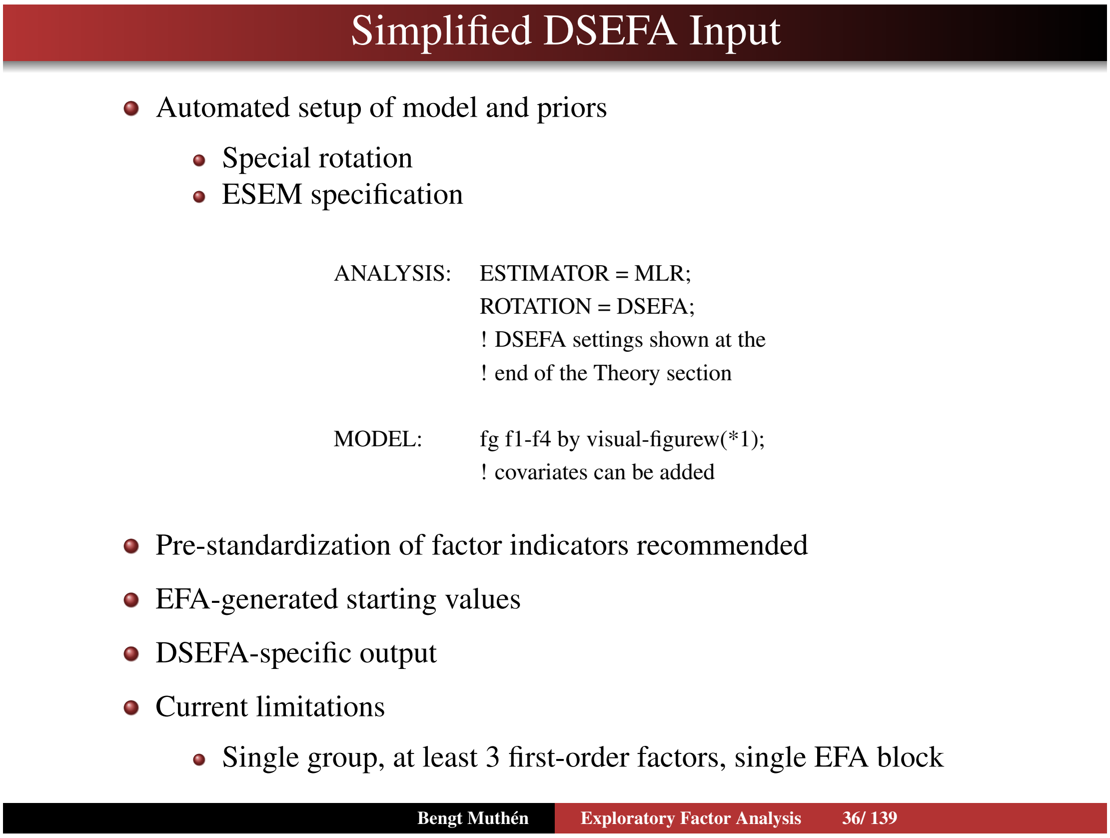
Input for Exploratory Bi-Factor Analysis Using DSEFA: Direct Effects Second-Order Factor Analysis

```
ANALYSIS:             ESTIMATOR = MLR;
                      ITERATIONS = 10000;
                      CONVERGENCE = 0.000001;
                      !STARTS = 50;

MODEL:                ! first 3 lines below same as SEFA:
                      f1-f4 BY visual-figurew*(a1-a76); ! 4*19=76
                      f1-f4 (v1-v4);
                      f0 BY f1-f4* (l1-l4);
                      ! direct effects from the second-order
                      ! factor to the factor indicators:
                      f0 BY visual-figurew*0 (d1-d19);
                      f0@1;

MODEL CONSTRAINT:     v1 = 1 - l1*l1;
                      v2 = 1 - l2*l2;
                      v3 = 1 - l3*l3;
                      v4 = 1 - l4*l4;

MODEL PRIORS:         a1-a76 ~ GEOMIN(4,0.1);
                      ! impose alignment function (ALF) priors:
                      d1-d19 ~ ALF(0,1);
```

Slide 35 shows the input for DSEFA. The ANALYSIS command includes settings for iterations, convergence, and starts that may be needed in the manual approach but not in the simplified, automatic approach given on the next slide.

In the MODEL command, the first 3 lines are the same as for SEFA. The specification of the second-order factor is bolded. After this, the direct effects from the second-order factor to the indicators are specified using the BY option (also bolded). Labels that are given for these effects (loadings) and used in the MODEL PRIORS command for the ALF priors.

The ALF priors have mean 0 and variance 1. Here, variance actually is 1/weight where the weight refers to the amount of penalty of the penalized maximum-likelihood estimation. The smaller the prior variance, the further the direct effects are pushed toward their means of zero. For a very small prior variance, it is as if there are no direct effects and the model is a SEFA. A variance of 1 is often a suitable choice. If necessary, it should be varied until the fit of the DSEFA is the same as for EFA with the same number of factors (here 5). For factor indicators that have widely varying variances, standardization of the indicators using the DEFINE command is recommended for the ALF priors to work well.



Simplified DSEFA Input

- Automated setup of model and priors
    - Special rotation
    - ESEM specification

```
ANALYSIS:   ESTIMATOR = MLR;
            ROTATION = DSEFA;
            ! DSEFA settings shown at the
            ! end of the Theory section

MODEL:      fg f1-f4 by visual-figurew(*1);
            ! covariates can be added
```

- Pre-standardization of factor indicators recommended
- EFA-generated starting values
- DSEFA-specific output
- Current limitations
    - Single group, at least 3 first-order factors, single EFA block

Slide 36 shows the simplified, automated DSEFA input. DSEFA is initiated by a special rotation called DSEFA.

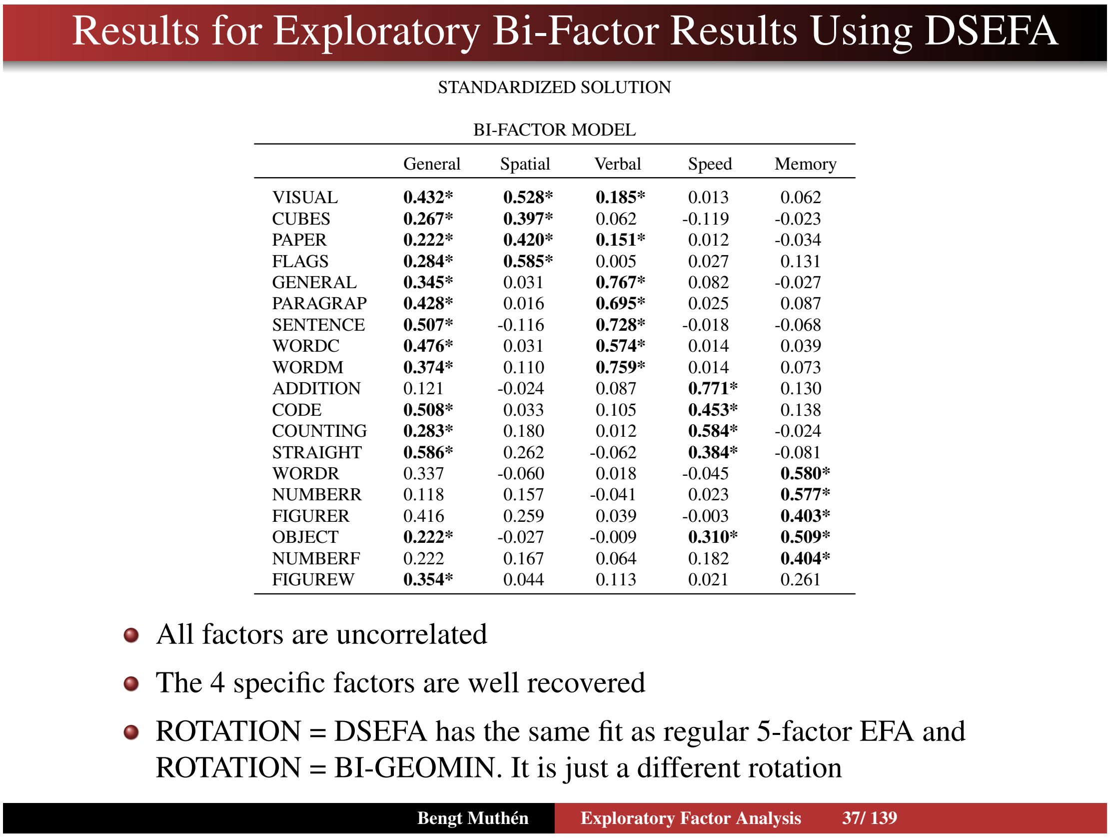
The MODEL command uses an ESEM style specification of the factors where the first factor is the general one.

Pre-standardization of the factor indicators is recommended.

EFA-generated starting values are provided so the special ANALYSIS settings mentioned on the previous slide are typically not needed.

The DSEFA output includes the total effects of the second-order factor on the indicators, computed as the sum of the indirect effects and the direct effects. See also the EFA Theory section on the SEFA - DSEFA transition.

Currently DSEFA is limited to 1 group. At least 3 first-order indicators are needed. And, there cannot be more than one EFA block.



Results for Exploratory Bi-Factor Results Using DSEFA

STANDARDIZED SOLUTION

BI-FACTOR MODEL

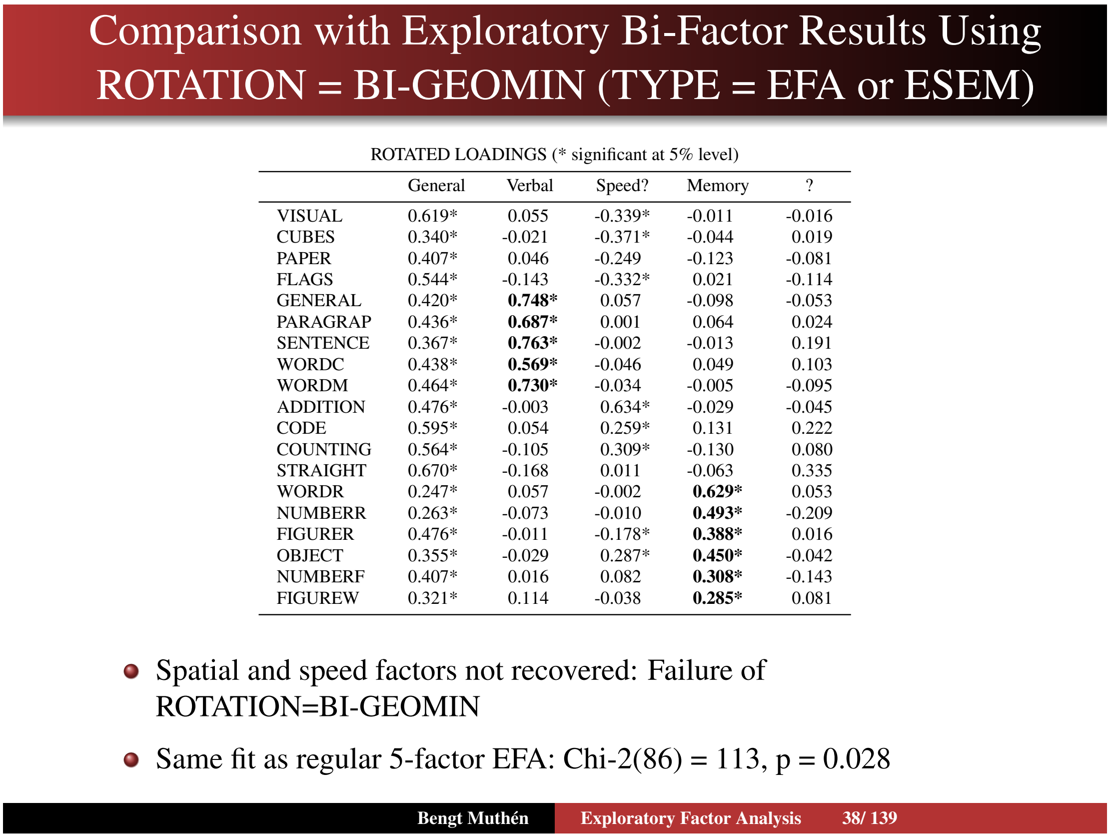
	General	Spatial	Verbal	Speed	Memory
VISUAL	**0.432***	**0.528***	**0.185***	0.013	0.062
CUBES	**0.267***	**0.397***	0.062	-0.119	-0.023
PAPER	**0.222***	**0.420***	**0.151***	0.012	-0.034
FLAGS	**0.284***	**0.585***	0.005	0.027	0.131
GENERAL	**0.345***	0.031	**0.767***	0.082	-0.027
PARAGRAP	**0.428***	0.016	**0.695***	0.025	0.087
SENTENCE	**0.507***	-0.116	**0.728***	-0.018	-0.068
WORDC	**0.476***	0.031	**0.574***	0.014	0.039
WORDM	**0.374***	0.110	**0.759***	0.014	0.073
ADDITION	0.121	-0.024	0.087	**0.771***	0.130
CODE	**0.508***	0.033	0.105	**0.453***	0.138
COUNTING	**0.283***	0.180	0.012	**0.584***	-0.024
STRAIGHT	**0.586***	0.262	-0.062	**0.384***	-0.081
WORDR	0.337	-0.060	0.018	-0.045	**0.580***
NUMBERR	0.118	0.157	-0.041	0.023	**0.577***
FIGURER	0.416	0.259	0.039	-0.003	**0.403***
OBJECT	**0.222***	-0.027	-0.009	**0.310***	**0.509***
NUMBERF	0.222	0.167	0.064	0.182	**0.404***
FIGUREW	**0.354***	0.044	0.113	0.021	0.261

- All factors are uncorrelated
- The 4 specific factors are well recovered
- ROTATION = DSEFA has the same fit as regular 5-factor EFA and ROTATION = BI-GEOMIN. It is just a different rotation

Slide 37 shows the DSEFA estimates. This output is presented under the headings STANDARDIZED SOLUTION, BI-FACTOR MODEL. All the factors are uncorrelated.

The general factor has mostly significant loadings. The 4 specific factors Spatial, Verbal, Speed, and Memory all have large significant loadings in expected places.

ROTATION = DSEFA has the same fit as regular 5-factor EFA and ROTATION = BI-GEOMIN. It is just a different rotation.



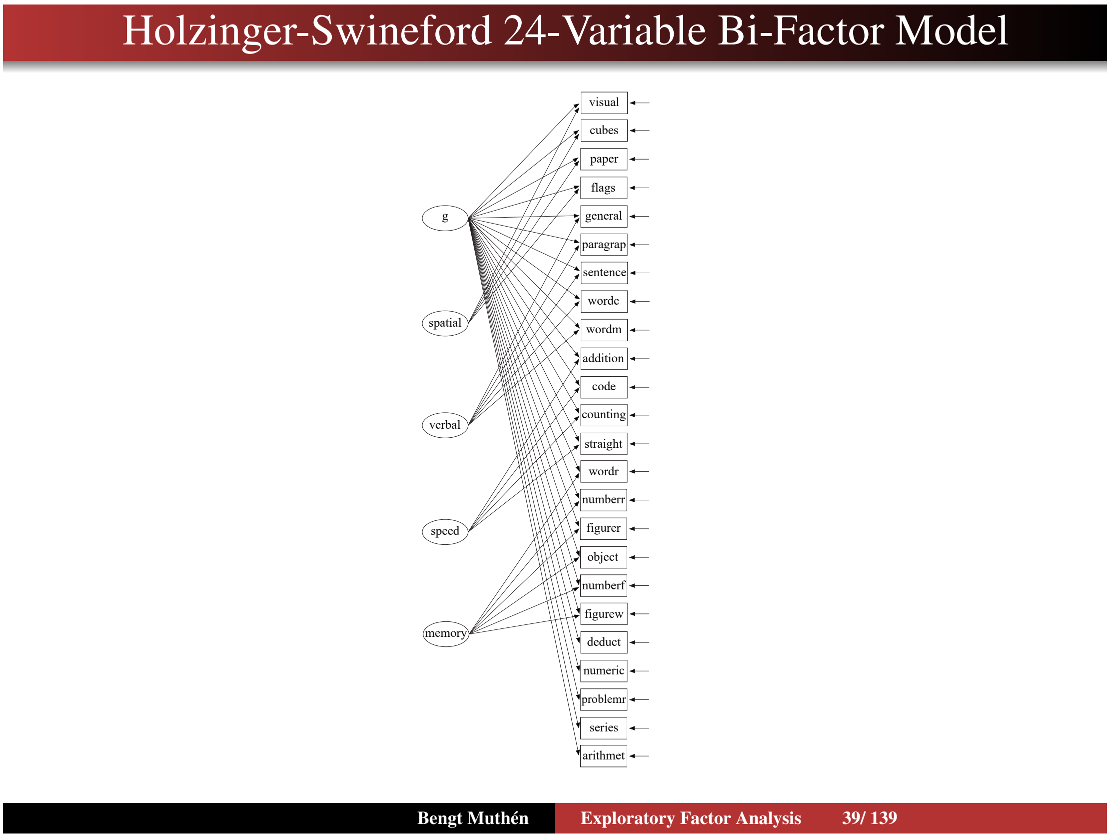
Comparison with Exploratory Bi-Factor Results Using ROTATION = BI-GEOMIN (TYPE = EFA or ESEM)

ROTATED LOADINGS (* significant at 5% level)

	General	Verbal	Speed?	Memory	?
VISUAL	0.619*	0.055	-0.339*	-0.011	-0.016
CUBES	0.340*	-0.021	-0.371*	-0.044	0.019
PAPER	0.407*	0.046	-0.249	-0.123	-0.081
FLAGS	0.544*	-0.143	-0.332*	0.021	-0.114
GENERAL	0.420*	**0.748***	0.057	-0.098	-0.053
PARAGRAP	0.436*	**0.687***	0.001	0.064	0.024
SENTENCE	0.367*	**0.763***	-0.002	-0.013	0.191
WORDC	0.438*	**0.569***	-0.046	0.049	0.103
WORDM	0.464*	**0.730***	-0.034	-0.005	-0.095
ADDITION	0.476*	-0.003	0.634*	-0.029	-0.045
CODE	0.595*	0.054	0.259*	0.131	0.222
COUNTING	0.564*	-0.105	0.309*	-0.130	0.080
STRAIGHT	0.670*	-0.168	0.011	-0.063	0.335
WORDR	0.247*	0.057	-0.002	**0.629***	0.053
NUMBERR	0.263*	-0.073	-0.010	**0.493***	-0.209
FIGURER	0.476*	-0.011	-0.178*	**0.388***	0.016
OBJECT	0.355*	-0.029	0.287*	**0.450***	-0.042
NUMBERF	0.407*	0.016	0.082	**0.308***	-0.143
FIGUREW	0.321*	0.114	-0.038	**0.285***	0.081

- Spatial and speed factors not recovered: Failure of ROTATION=BI-GEOMIN
- Same fit as regular 5-factor EFA: Chi-2(86) = 113, p = 0.028

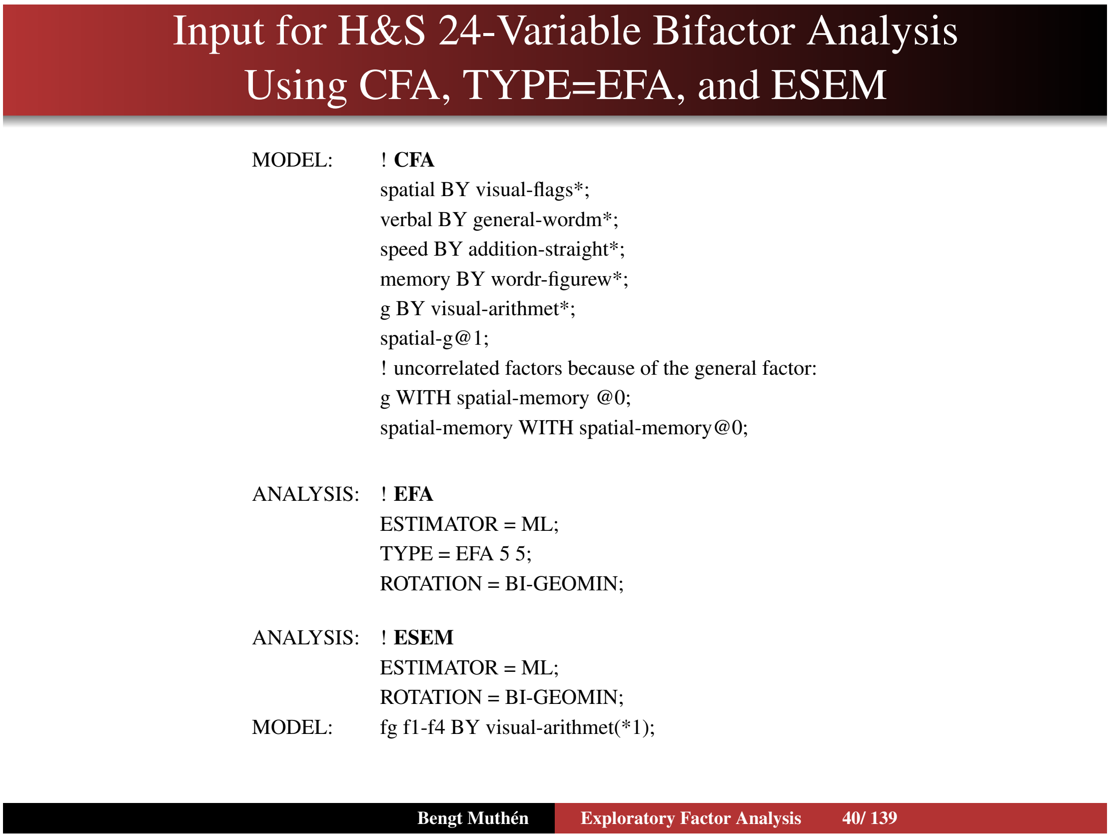
Slide 38 uses the older approach of ROTATION = BI-GEOMIN (Jennrich & Bentler, 2011, 2012). Although the model fit for ROTATION = BI-GEOMIN is the same as for ROTATION = DSEFA on the previous slide, the solution using ROTATION = BI-GEOMIN does not give the same quality of results. With this rotation, only the Verbal and Memory factors are well recovered, whereas the Spatial factor is not recovered and the Speed factor is poorly recovered.



Holzinger-Swineford 24-Variable Bi-Factor Model

Bengt Muthén Exploratory Factor Analysis 39/ 139

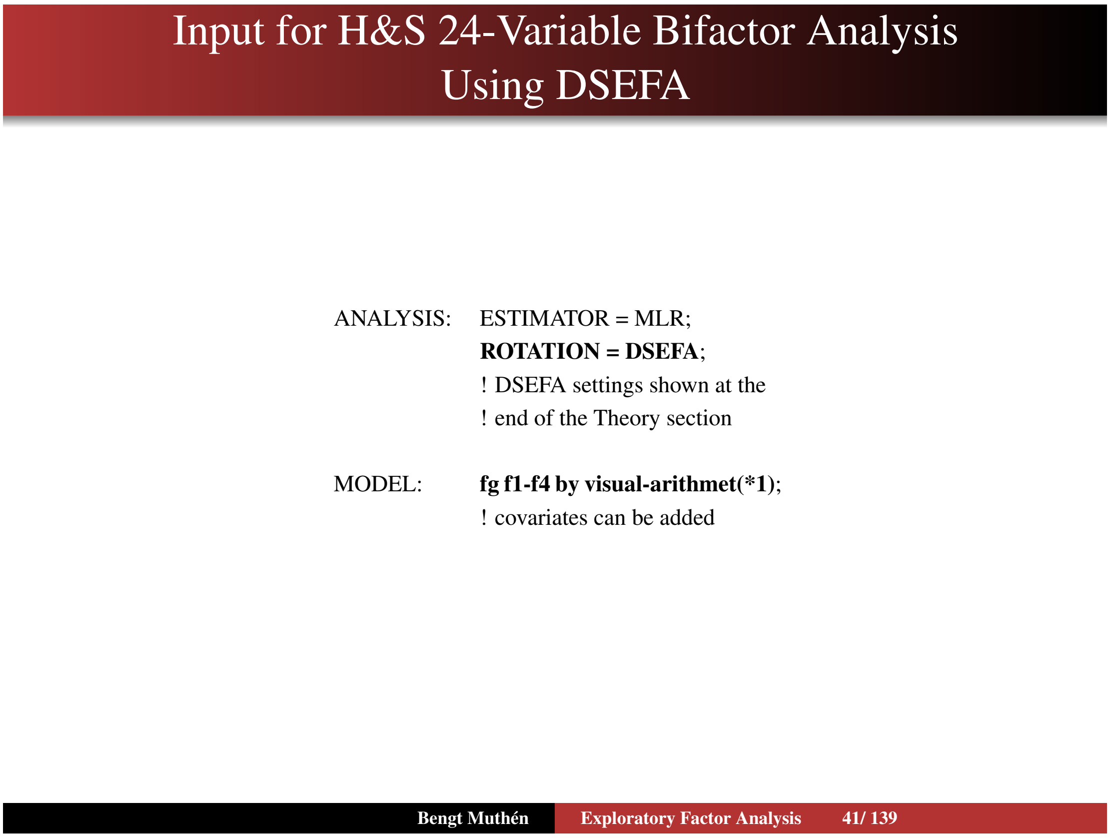
Slide 39 shows a bi-factor model diagram for the H&S data extended by 5 tests to give 24 factor indicators. The 5 test are at the bottom of the diagram. The diagram shows them loading only on the general factor.



Input for H&S 24-Variable Bifactor Analysis Using CFA, TYPE=EFA, and ESEM

```
MODEL:      ! CFA
            spatial BY visual-flags*;
            verbal BY general-wordm*;
            speed BY addition-straight*;
            memory BY wordr-figurew*;
            g BY visual-arithmet*;
            spatial-g@1;
            ! uncorrelated factors because of the general factor:
            g WITH spatial-memory @0;
            spatial-memory WITH spatial-memory@0;

ANALYSIS:   ! EFA
            ESTIMATOR = ML;
            TYPE = EFA 5 5;
            ROTATION = BI-GEOMIN;

ANALYSIS:   ! ESEM
            ESTIMATOR = ML;
            ROTATION = BI-GEOMIN;
MODEL:      fg f1-f4 BY visual-arithmet(*1);
```

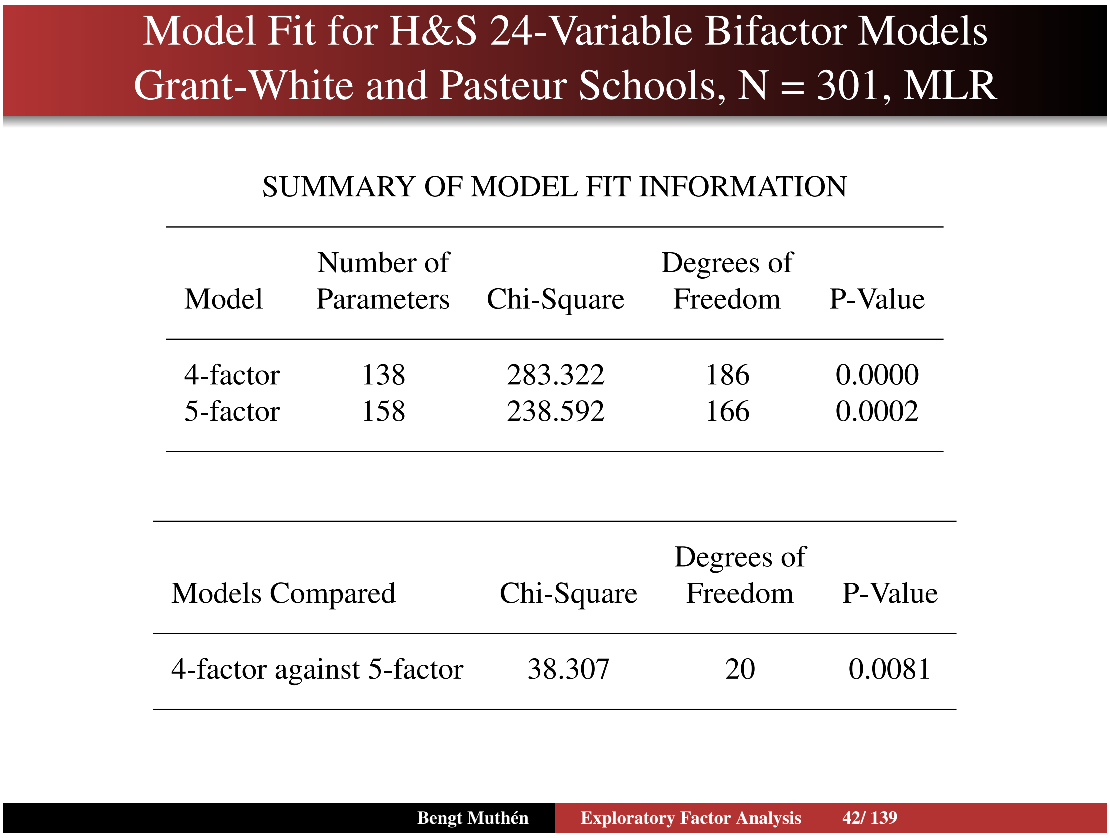
Slide 40 shows 3 inputs for these data: CFA, EFA, and ESEM. The EFA and ESEM results using ROTATION = BI-GEOMIN will be compared to DSEFA.



Input for H&S 24-Variable Bifactor Analysis Using DSEFA

```
ANALYSIS:   ESTIMATOR = MLR;
            ROTATION = DSEFA;
            ! DSEFA settings shown at the
            ! end of the Theory section

MODEL:      fg f1-f4 by visual-arithmet(*1);
            ! covariates can be added
```

Slide 41 shows the DSEFA input for the 24-variable analysis. DSEFA is defined by ROTATION = DSEFA. The DSEFA option can have several settings related to the priors. These are described in the EFA Theory section. The MODEL statement uses an ESEM type specification.



Model Fit for H&S 24-Variable Bifactor Models
Grant-White and Pasteur Schools, N = 301, MLR

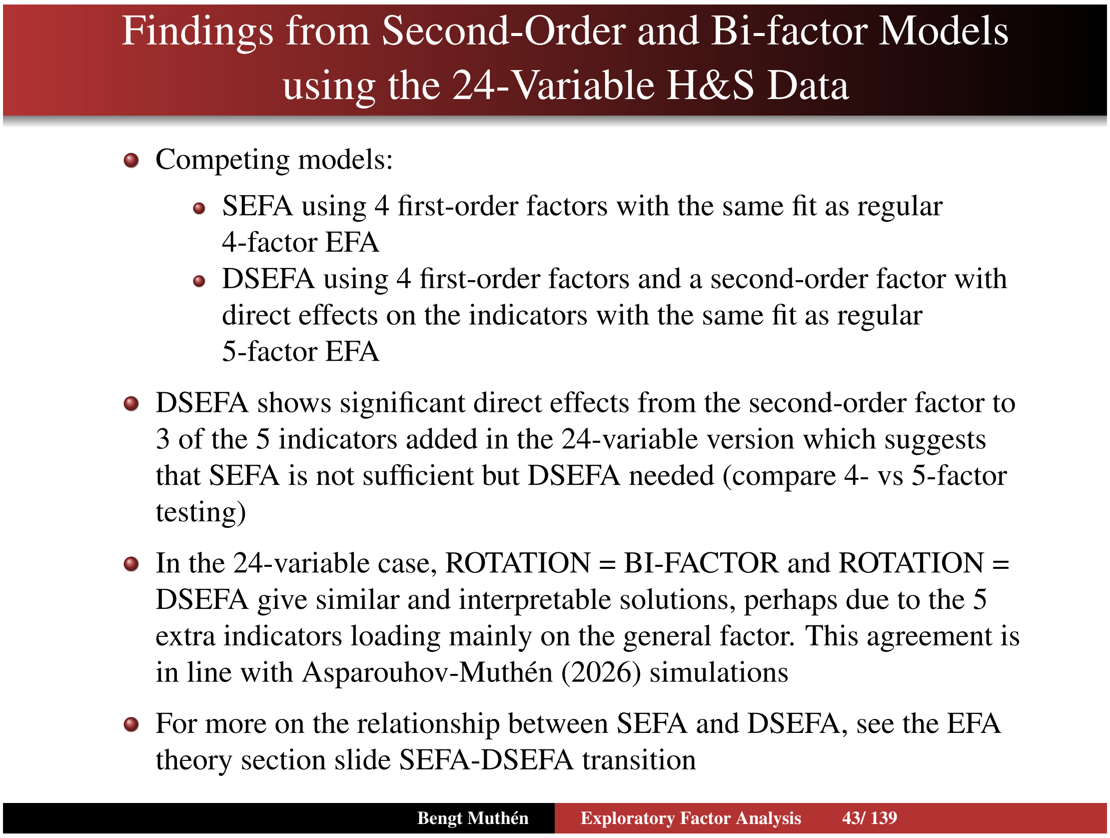
SUMMARY OF MODEL FIT INFORMATION

Model	Number of Parameters	Chi-Square	Degrees of Freedom	P-Value
4-factor	138	283.322	186	0.0000
5-factor	158	238.592	166	0.0002

Models Compared	Chi-Square	Degrees of Freedom	P-Value
4-factor against 5-factor	38.307	20	0.0081

Slide 42 shows EFA model test results for the 24-variable analysis. In contrast to the 19-variable case, the fit is not good for 4 factors and the 5-factor solution is chosen here. For the bi-factor analyses, a 5-factor solution corresponds to one general and 4 specific factors.

The fit of the 5-factor model can be improved by searching for significant residual covariances. With PSEM, modification indices are not available to guide this search. A simple approach is to instead use an ESEM EFA analysis with 5 factors requesting modindices since the fit of this model is the same as for the DSEFA. This pointed to a significant residual covariance between figurew and addition. When this was freed in the DSEFA analysis, a better model fit was obtained: chi-square(165) = 202 (p = 0.025).



Findings from Second-Order and Bi-factor Models using the 24-Variable H&S Data

- Competing models:
  - SEFA using 4 first-order factors with the same fit as regular 4-factor EFA
  - DSEFA using 4 first-order factors and a second-order factor with direct effects on the indicators with the same fit as regular 5-factor EFA
- DSEFA shows significant direct effects from the second-order factor to 3 of the 5 indicators added in the 24-variable version which suggests that SEFA is not sufficient but DSEFA needed (compare 4- vs 5-factor testing)
- In the 24-variable case, ROTATION = BI-FACTOR and ROTATION = DSEFA give similar and interpretable solutions, perhaps due to the 5 extra indicators loading mainly on the general factor. This agreement is in line with Asparouhov-Muthén (2026) simulations
- For more on the relationship between SEFA and DSEFA, see the EFA theory section slide SEFA-DSEFA transition

Bengt Muthén **Exploratory Factor Analysis** **43/ 139**

Slide 43 discusses results from second-order and bi-factor analysis for the 24-variable H&S data.

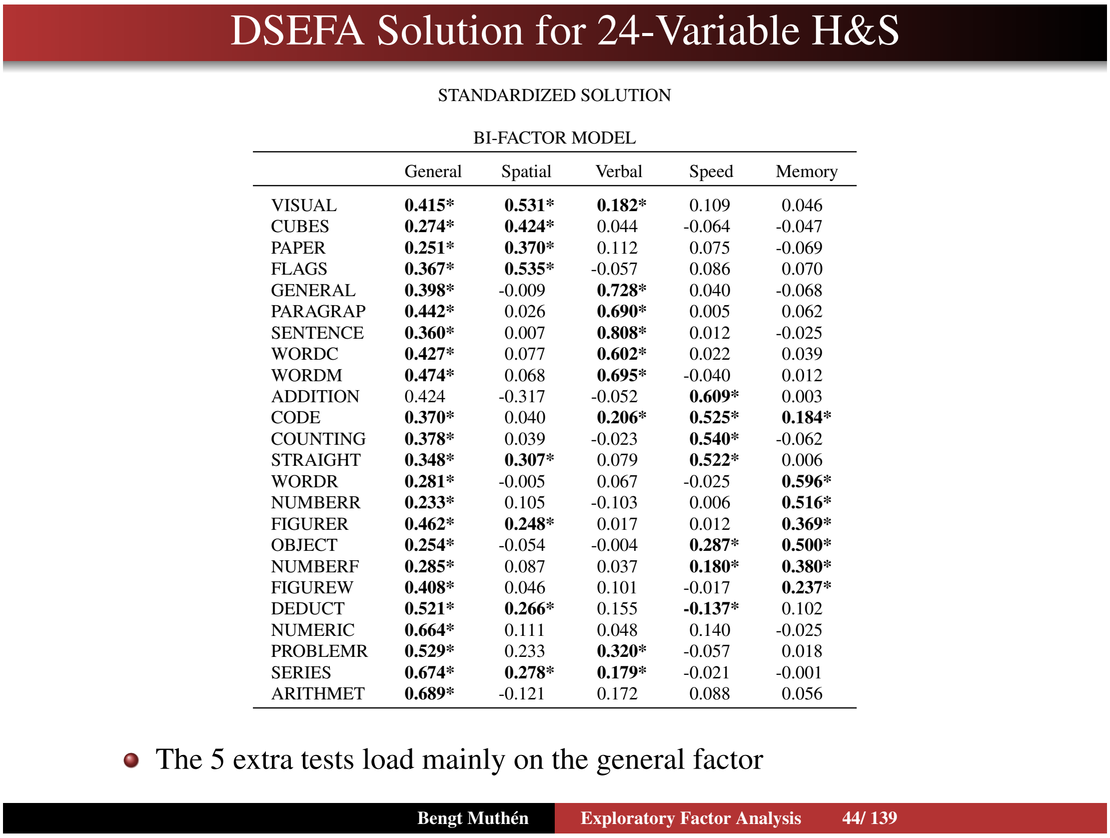
We consider two competing models:

- SEFA using 4 first-order factors with the same fit as regular 4-factor EFA
- DSEFA using 4 first-order factors and a second-order factor with direct effects on the indicators with the same fit as regular 5-factor EFA

DSEFA shows significant direct effects from the second-order factor to 3 of the 5 indicators added in the 24-variable version, which suggests that SEFA is not sufficient but DSEFA is needed (compare 4- vs 5-factor EFA testing).

In the 24-variable case, ROTATION = BI-FACTOR and ROTATION = DSEFA give similar and interpretable solutions, perhaps due to the 5 extra indicators loading mainly on the general factor. The agreement in this case is in line with Asparouhov-Muthén (2026) simulations.

Slide 135 in the EFA theory section says a bit about the interpretation of the different factor loading matrices for SEFA and DSEFA and how the models relate to each other.



DSEFA Solution for 24-Variable H&S

STANDARDIZED SOLUTION

BI-FACTOR MODEL

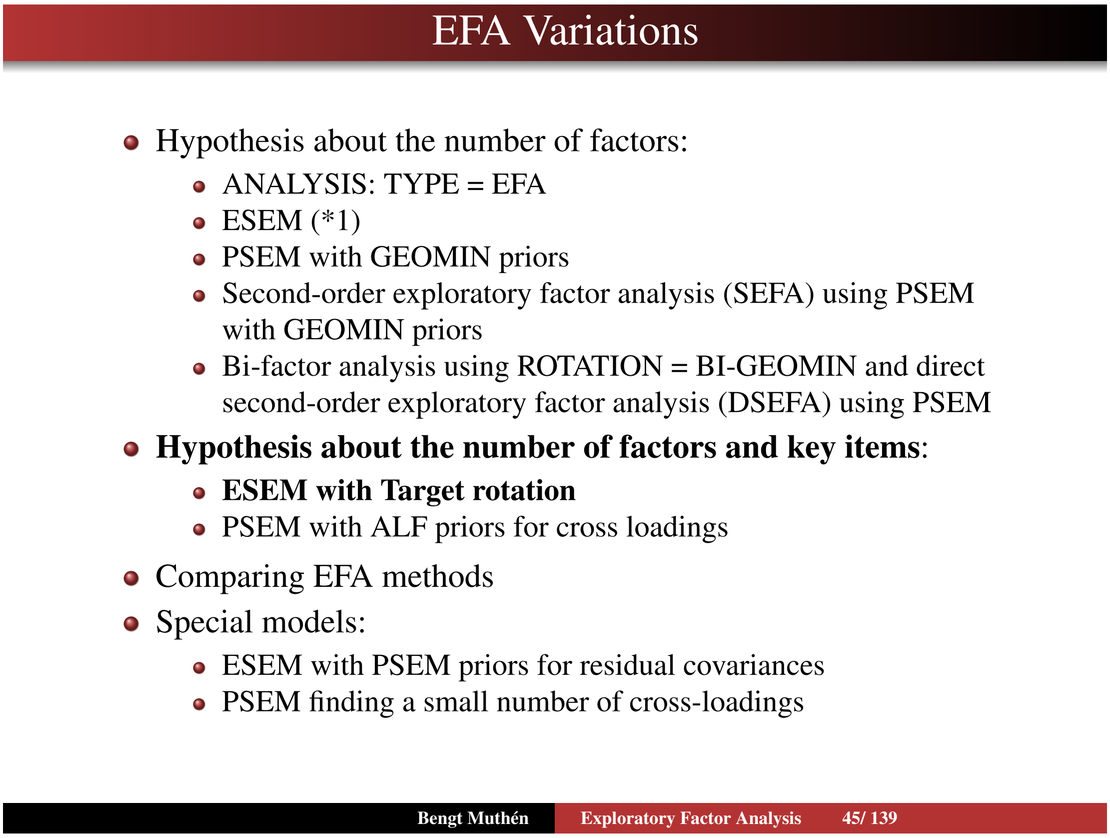
	General	Spatial	Verbal	Speed	Memory
VISUAL	**0.415***	**0.531***	**0.182***	0.109	0.046
CUBES	**0.274***	**0.424***	0.044	-0.064	-0.047
PAPER	**0.251***	**0.370***	0.112	0.075	-0.069
FLAGS	**0.367***	**0.535***	-0.057	0.086	0.070
GENERAL	**0.398***	-0.009	**0.728***	0.040	-0.068
PARAGRAP	**0.442***	0.026	**0.690***	0.005	0.062
SENTENCE	**0.360***	0.007	**0.808***	0.012	-0.025
WORDC	**0.427***	0.077	**0.602***	0.022	0.039
WORDM	**0.474***	0.068	**0.695***	-0.040	0.012
ADDITION	0.424	-0.317	-0.052	**0.609***	0.003
CODE	**0.370***	0.040	**0.206***	**0.525***	**0.184***
COUNTING	**0.378***	0.039	-0.023	**0.540***	-0.062
STRAIGHT	**0.348***	**0.307***	0.079	**0.522***	0.006
WORDR	**0.281***	-0.005	0.067	-0.025	**0.596***
NUMBERR	**0.233***	0.105	-0.103	0.006	**0.516***
FIGURER	**0.462***	**0.248***	0.017	0.012	**0.369***
OBJECT	**0.254***	-0.054	-0.004	**0.287***	**0.500***
NUMBERF	**0.285***	0.087	0.037	**0.180***	**0.380***
FIGUREW	**0.408***	0.046	0.101	-0.017	**0.237***
DEDUCT	**0.521***	**0.266***	0.155	**-0.137***	0.102
NUMERIC	**0.664***	0.111	0.048	0.140	-0.025
PROBLEMR	**0.529***	0.233	**0.320***	-0.057	0.018
SERIES	**0.674***	**0.278***	**0.179***	-0.021	-0.001
ARITHMET	**0.689***	-0.121	0.172	0.088	0.056

- The 5 extra tests load mainly on the general factor

Slide 44 shows the DSEFA solution for the 24-variable case. This output is presented under the headings STANDARDIZED SOLUTION, BI-FACTOR MODEL. All the factors are uncorrelated.

The 4 specific factors are all well recovered. The general factor has significant loadings for all variables. The 5 extra variables mainly load on the general factor with some cross loadings.

2.2 Hypothesis about the number of factors and key items



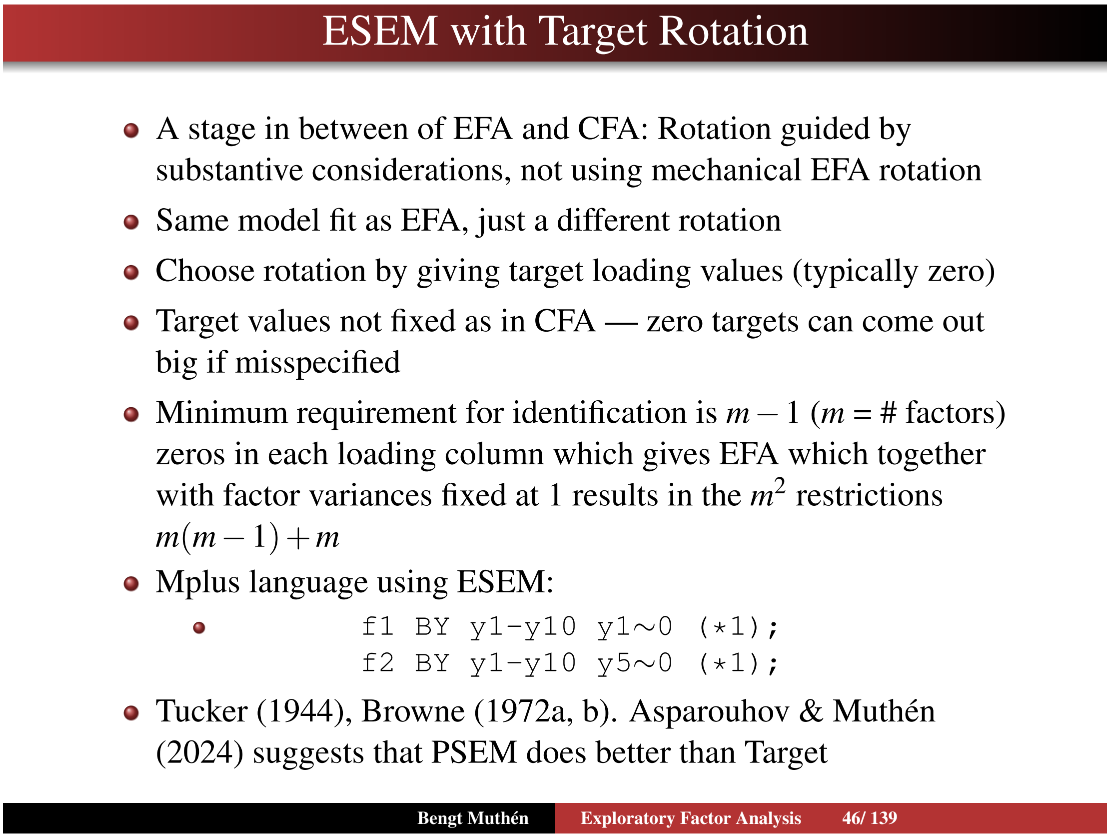
EFA Variations

- Hypothesis about the number of factors:
    - ANALYSIS: TYPE = EFA
    - ESEM (*1)
    - PSEM with GEOMIN priors
    - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
    - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- **Hypothesis about the number of factors and key items**:
    - **ESEM with Target rotation**
    - PSEM with ALF priors for cross loadings
- Comparing EFA methods
- Special models:
    - ESEM with PSEM priors for residual covariances
    - PSEM finding a small number of cross-loadings

Slide 45 returns to the EFA Variations overview where we now turn to analyses where not only the number of factors is specified but also key items. This was referred to as the middle analysis stage on slide 4.

The first topic is ESEM with target rotation.

2.2.1 ESEM with Target rotation



ESEM with Target Rotation

- A stage in between of EFA and CFA: Rotation guided by substantive considerations, not using mechanical EFA rotation
- Same model fit as EFA, just a different rotation
- Choose rotation by giving target loading values (typically zero)
- Target values not fixed as in CFA — zero targets can come out big if misspecified
- Minimum requirement for identification is $m-1$ ($m = \#$ factors) zeros in each loading column which gives EFA which together with factor variances fixed at 1 results in the m^2 restrictions $m(m-1)+m$
- Mplus language using ESEM:

  -
    ```
    f1 BY y1-y10 y1~0 (*1);
    f2 BY y1-y10 y5~0 (*1);
    ```
- Tucker (1944), Browne (1972a, b). Asparouhov & Muthén (2024) suggests that PSEM does better than Target

Slide 46 discusses Target rotation. Target rotation represents an analysis stage that lies in between of EFA and CFA: The rotation is guided by substantive considerations, not using mechanical EFA rotation.

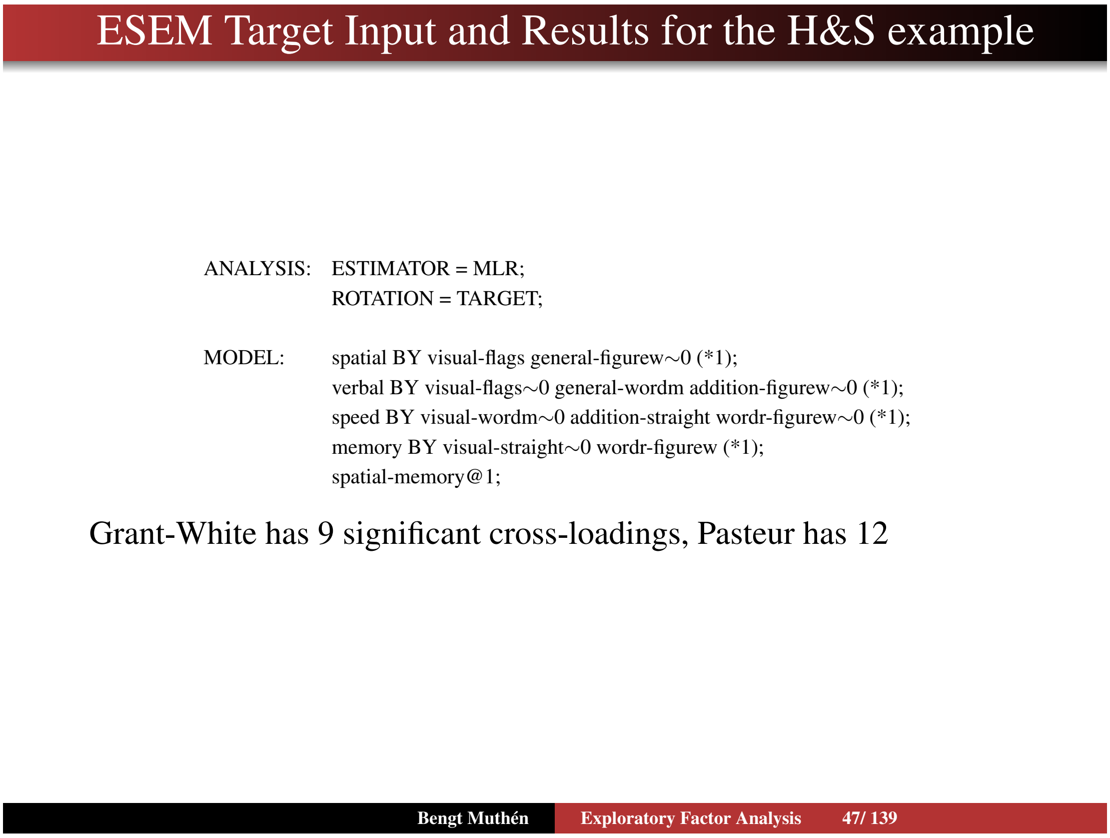
It has the same model fit as EFA, just using a different rotation. You choose a rotation by giving target loading values (typically zero).

Target values are not fixed as in CFA — zero targets can come out big if misspecified.

Minimum requirement for identification is $m-1$ (where $m = \#$factors) zeros in each loading column which gives EFA which together with factor variances fixed at 1 results in the $m(m-1)+m = m^2$ restrictions.

The Mplus language uses ESEM for one factor at a time using the label 1 to denote the same EFA block. A curl $\sim$ is used together with a value to denote a target value for a loading, typically zero.

At the bottom are two early references for this technique and also the 2024 PSEM paper by us. As we will see, there are now better approaches to target rotation.



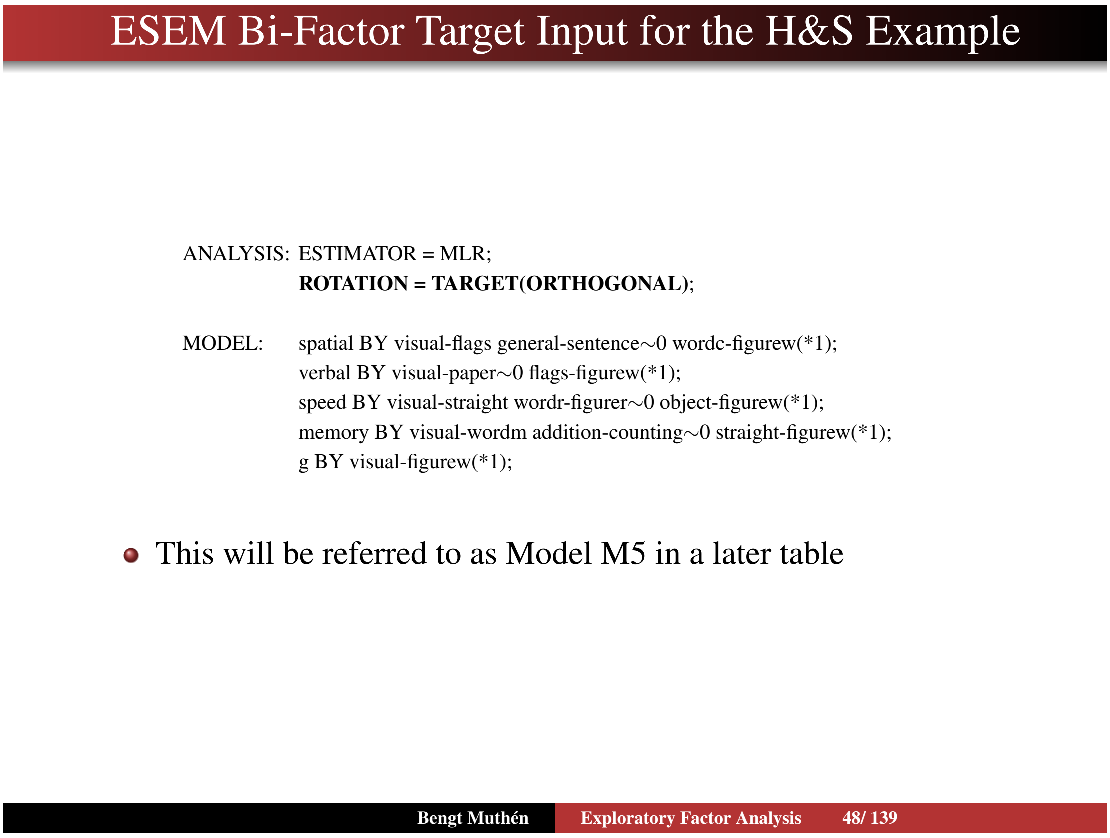
ESEM Target Input and Results for the H&S example

```
ANALYSIS:   ESTIMATOR = MLR;
            ROTATION = TARGET;

MODEL:      spatial BY visual-flags general-figurew~0 (*1);
            verbal BY visual-flags~0 general-wordm addition-figurew~0 (*1);
            speed BY visual-wordm~0 addition-straight wordr-figurew~0 (*1);
            memory BY visual-straight~0 wordr-figurew (*1);
            spatial-memory@1;
```

Grant-White has 9 significant cross-loadings, Pasteur has 12

Slide 47 shows the target input for the H&S example with 19 variables. The ANALYSIS uses ROTATION = TARGET. The MODEL command uses the loading pattern of slide 9 to give zero target values to all cross loadings.

This analysis results in 9 significant cross loadings for the Grant- White school and 12 for Pasteur. This is more than the 5 and 8 that we saw with regular EFA.



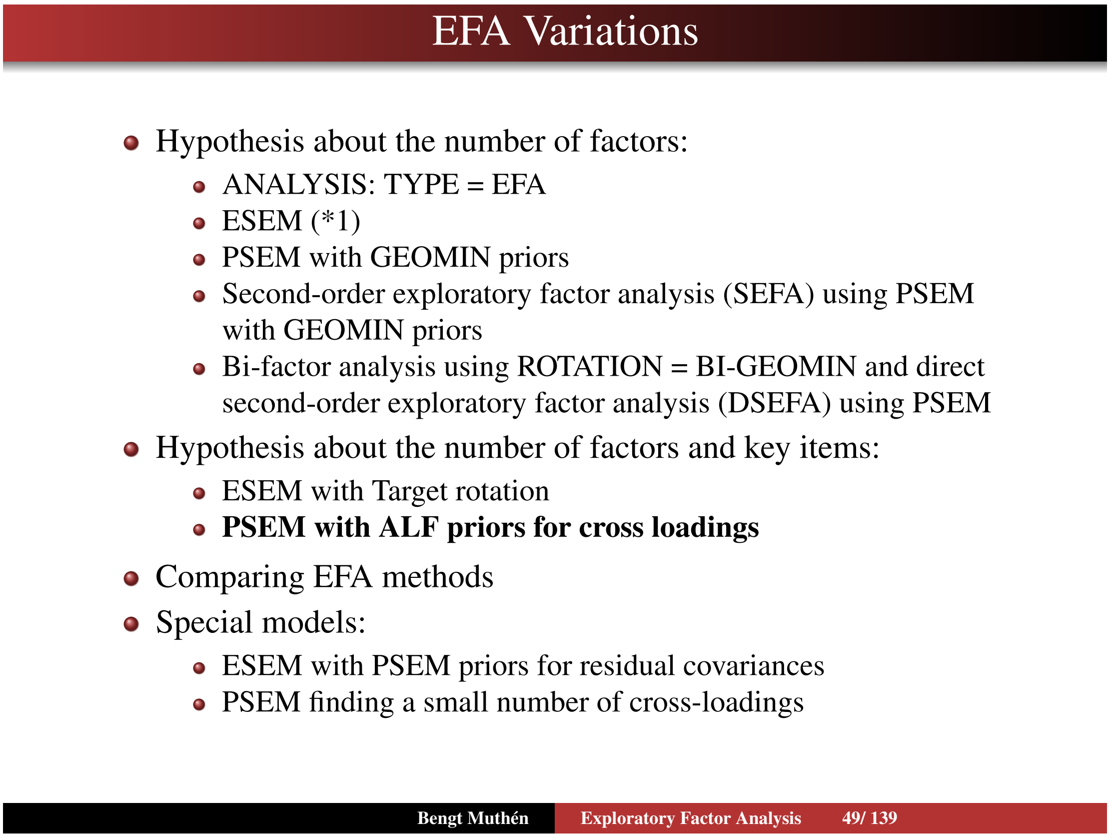
ESEM Bi-Factor Target Input for the H&S Example

```
ANALYSIS: ESTIMATOR = MLR;
          ROTATION = TARGET(ORTHOGONAL);

MODEL:    spatial BY visual-flags general-sentence~0 wordc-figurew(*1);
          verbal BY visual-paper~0 flags-figurew(*1);
          speed BY visual-straight wordr-figurer~0 object-figurew(*1);
          memory BY visual-wordm addition-counting~0 straight-figurew(*1);
          g BY visual-figurew(*1);
```

- This will be referred to as Model M5 in a later table

Slide 48 shows that Mplus also offers target rotation with bi-factor modeling. In the ANALYSIS command, this uses ROTATION = TARGET(ORTHOGONAL) so that all factors are uncorrelated. The MODEL command adds the general factor called g to the ESEM specification. On slides 56 and 57 this is the specification referred to as M5.

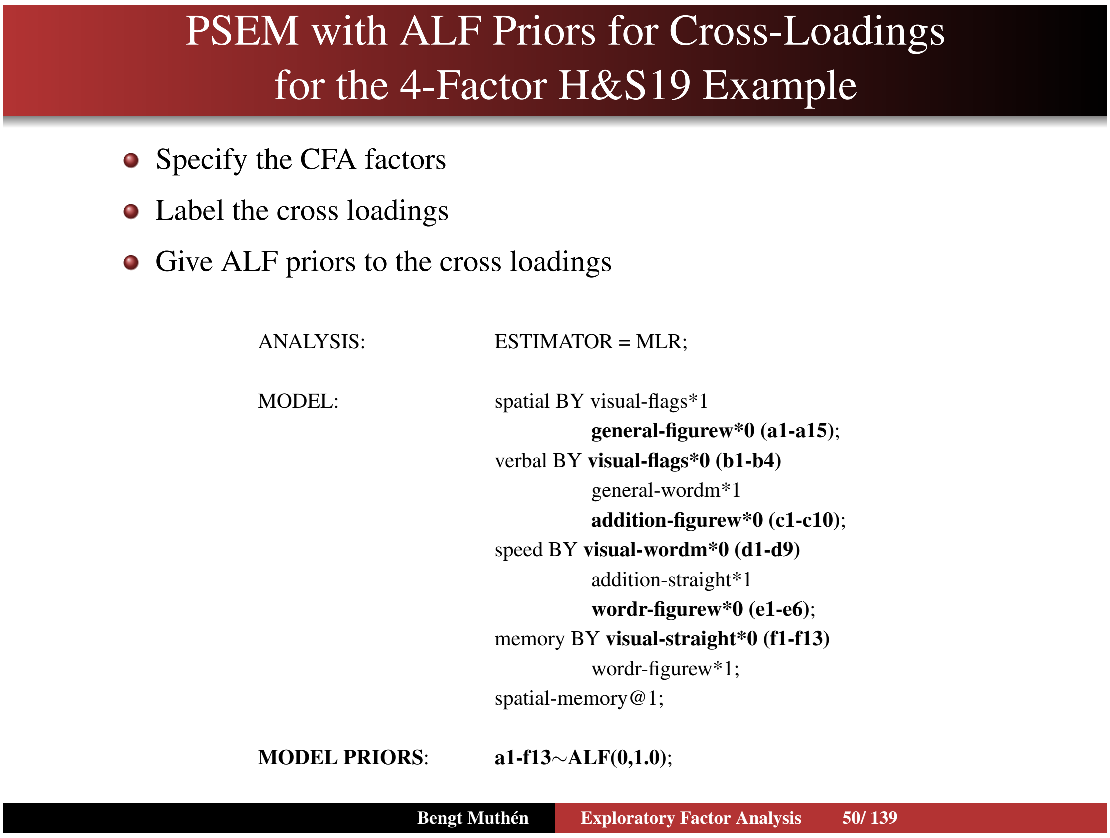
2.2.2 PSEM with ALF priors for cross loadings



EFA Variations

- Hypothesis about the number of factors:
  - ANALYSIS: TYPE = EFA
  - ESEM (*1)
  - PSEM with GEOMIN priors
  - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
  - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- Hypothesis about the number of factors and key items:
  - ESEM with Target rotation
  - **PSEM with ALF priors for cross loadings**
- Comparing EFA methods
- Special models:
  - ESEM with PSEM priors for residual covariances
  - PSEM finding a small number of cross-loadings

Slide 49 shows the EFA variations overview again and we see that we will now turn to using PSEM with ALF priors for cross loadings.



PSEM with ALF Priors for Cross-Loadings for the 4-Factor H&S19 Example

- Specify the CFA factors
- Label the cross loadings
- Give ALF priors to the cross loadings

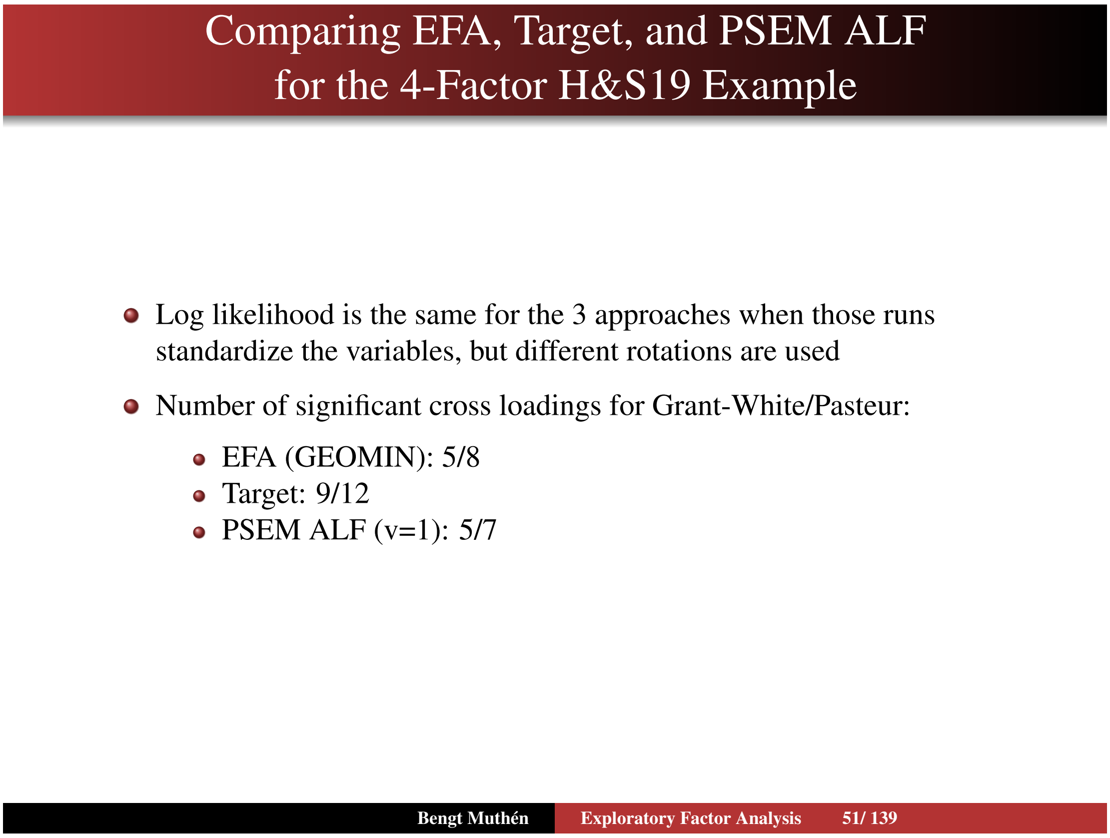
```
ANALYSIS:          ESTIMATOR = MLR;

MODEL:             spatial BY visual-flags*1
                         general-figurew*0 (a1-a15);
                   verbal BY visual-flags*0 (b1-b4)
                         general-wordm*1
                         addition-figurew*0 (c1-c10);
                   speed BY visual-wordm*0 (d1-d9)
                         addition-straight*1
                         wordr-figurew*0 (e1-e6);
                   memory BY visual-straight*0 (f1-f13)
                         wordr-figurew*1;
                   spatial-memory@1;

MODEL PRIORS:      a1-f13~ALF(0,1.0);
```

Slide 50 shows the input for PSEM with ALF priors for cross loadings. There are 3 parts to this: Specify the CFA factor loadings, label the cross loadings, and give ALF priors for the cross loadings in MODEL PRIORS.

The cross loadings are the same loadings that were given zero targets in the Target rotation approach. With ALF priors, the targets need not be exactly zero but approximately so by varying the variance of the ALF prior. Target rotation would correspond to using a very small ALF prior variance so that the values are essentially held at the prior mean of zero. The ALF prior variances can also be different for different cross loadings.

The intention of the PSEM ALF cross-loadings analysis is that its log likelihood value should match that of the 4-factor EFA so that PSEM ALF essentially provides a different rotation. To accomplish this, a variance of 1 is often a suitable choice for the ALF prior variance but this should be checked.

For factor indicators that have widely varying variances, standardization of the indicators using the DEFINE command is recommended for the ALF priors to work well.

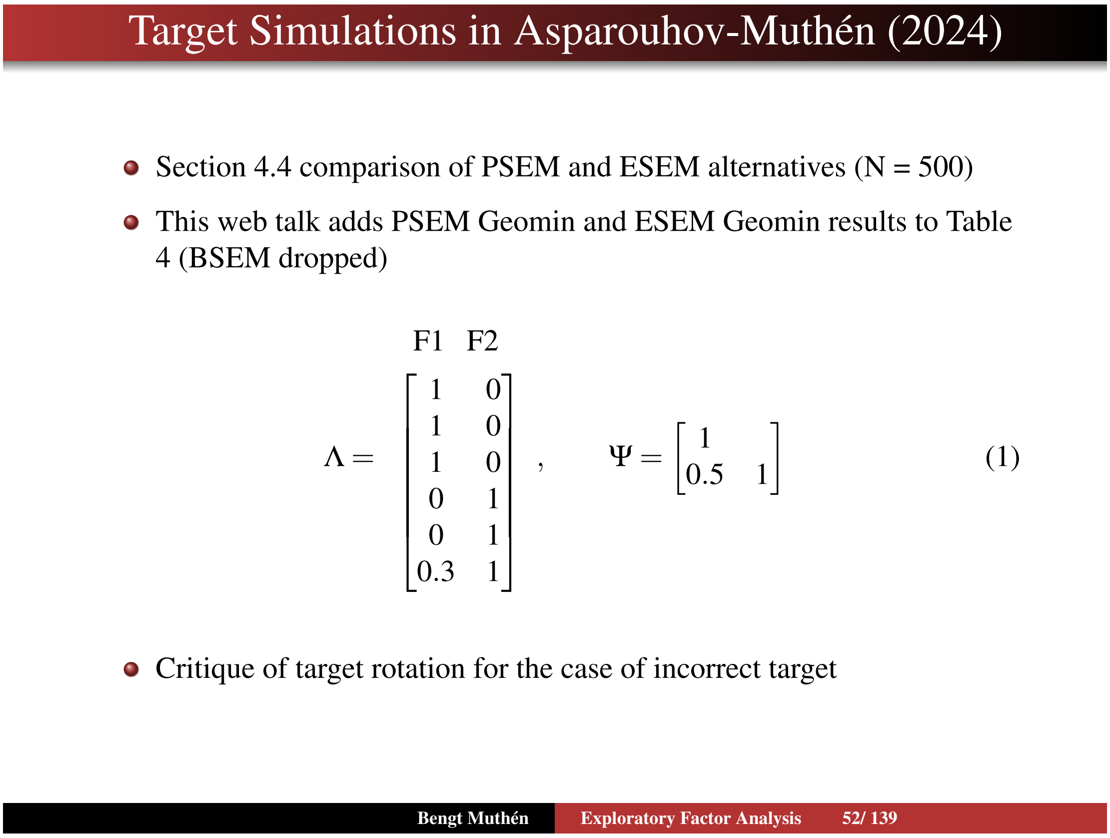


Comparing EFA, Target, and PSEM ALF for the 4-Factor H&S19 Example

- Log likelihood is the same for the 3 approaches when those runs standardize the variables, but different rotations are used
- Number of significant cross loadings for Grant-White/Pasteur:
    - EFA (GEOMIN): 5/8
    - Target: 9/12
    - PSEM ALF (v=1): 5/7

Slide 51 gives a comparison of EFA, Target rotation, and PSEM ALF cross loadings for the H&S example with 19 variables. The model fit is the same - the log likelihoods agree when the variables are standardized before analysis.

But, the 3 methods use different rotations producing different numbers of significant cross loadings for Grant-White/Pasteur: 5/8, 9/12, 5/7. The PSEM ALF cross loading approach appears to give the simplest solution.



Target Simulations in Asparouhov-Muthén (2024)

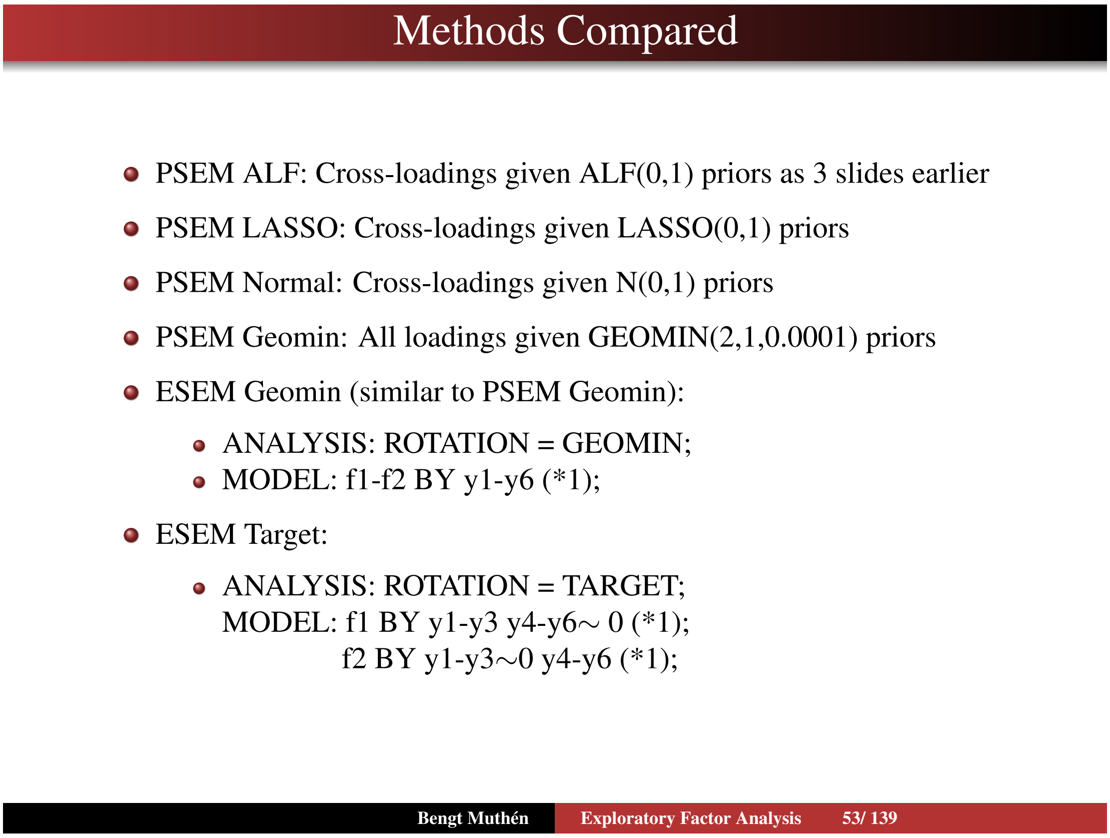
- Section 4.4 comparison of PSEM and ESEM alternatives (N = 500)
- This web talk adds PSEM Geomin and ESEM Geomin results to Table 4 (BSEM dropped)

$$\Lambda = \begin{array}{c} \begin{array}{cc} \text{F1} & \text{F2} \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0.3 & 1 \end{bmatrix} \end{array}, \qquad \Psi = \begin{bmatrix} 1 & \\ 0.5 & 1 \end{bmatrix} \tag{1}$$

- Critique of target rotation for the case of incorrect target

Slide 52 describes a simulation study in our 2024 PSEM paper. The simulations support that using PSEM ALF priors for cross loadings is preferred over Target rotation.

The model used for the simulations is shown in equation (1). The simple loading pattern is broken by one cross loading that appears for the last variable and the first factor. The premise is that the Target rotation incorrectly specified a zero target for this loading.



Methods Compared

- PSEM ALF: Cross-loadings given ALF(0,1) priors as 3 slides earlier
- PSEM LASSO: Cross-loadings given LASSO(0,1) priors
- PSEM Normal: Cross-loadings given N(0,1) priors
- PSEM Geomin: All loadings given GEOMIN(2,1,0.0001) priors
- ESEM Geomin (similar to PSEM Geomin):
  - ANALYSIS: ROTATION = GEOMIN;
  - MODEL: f1-f2 BY y1-y6 (*1);
- ESEM Target:
  - ANALYSIS: ROTATION = TARGET;
    MODEL: f1 BY y1-y3 y4-y6~ 0 (*1);
    f2 BY y1-y3~0 y4-y6 (*1);

Slide 53 lists the different approaches that are studied. The preferred PSEM ALF approach is at the top. The target rotation approach is listed last showing the incorrect zero target for the y6 loading of the f1 factor.

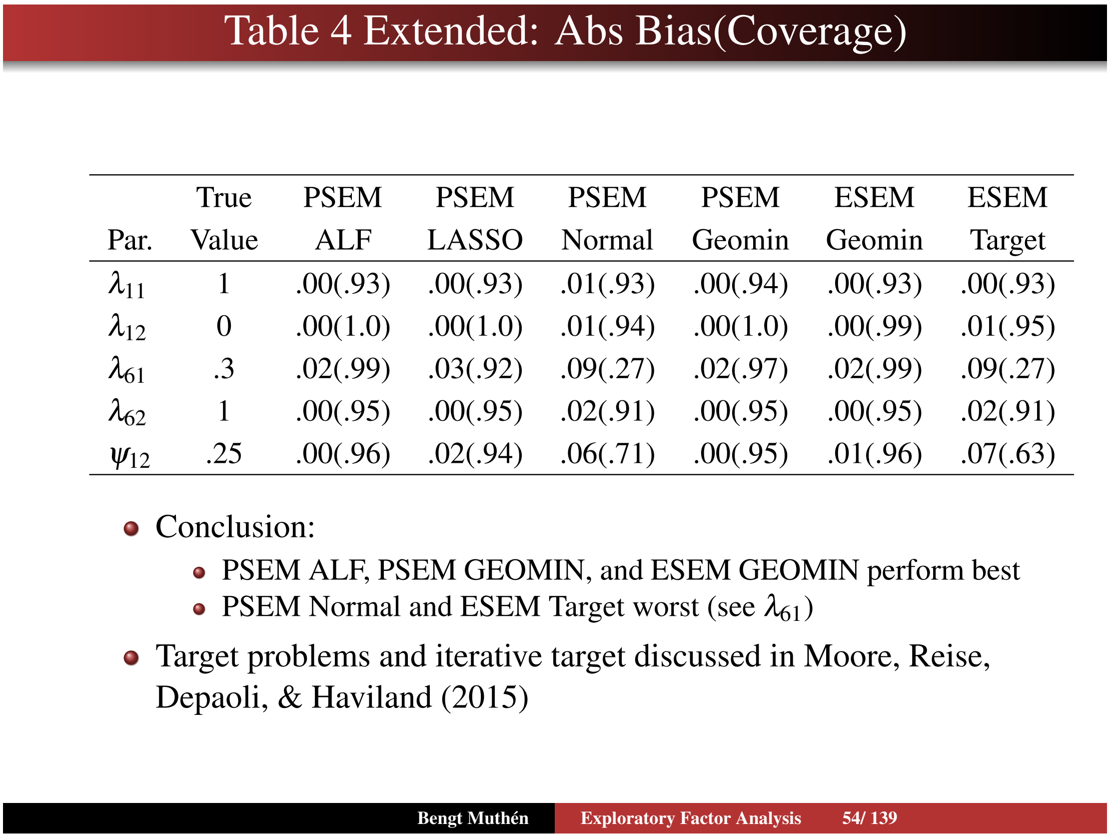


Table 4 Extended: Abs Bias(Coverage)

Par.	True Value	PSEM ALF	PSEM LASSO	PSEM Normal	PSEM Geomin	ESEM Geomin	ESEM Target
λ_{11}	1	.00(.93)	.00(.93)	.01(.93)	.00(.94)	.00(.93)	.00(.93)
λ_{12}	0	.00(1.0)	.00(1.0)	.01(.94)	.00(1.0)	.00(.99)	.01(.95)
λ_{61}	.3	.02(.99)	.03(.92)	.09(.27)	.02(.97)	.02(.99)	.09(.27)
λ_{62}	1	.00(.95)	.00(.95)	.02(.91)	.00(.95)	.00(.95)	.02(.91)
ψ_{12}	.25	.00(.96)	.02(.94)	.06(.71)	.00(.95)	.01(.96)	.07(.63)

- Conclusion:
  - PSEM ALF, PSEM GEOMIN, and ESEM GEOMIN perform best
  - PSEM Normal and ESEM Target worst (see λ_{61})
- Target problems and iterative target discussed in Moore, Reise, Depaoli, & Haviland (2015)

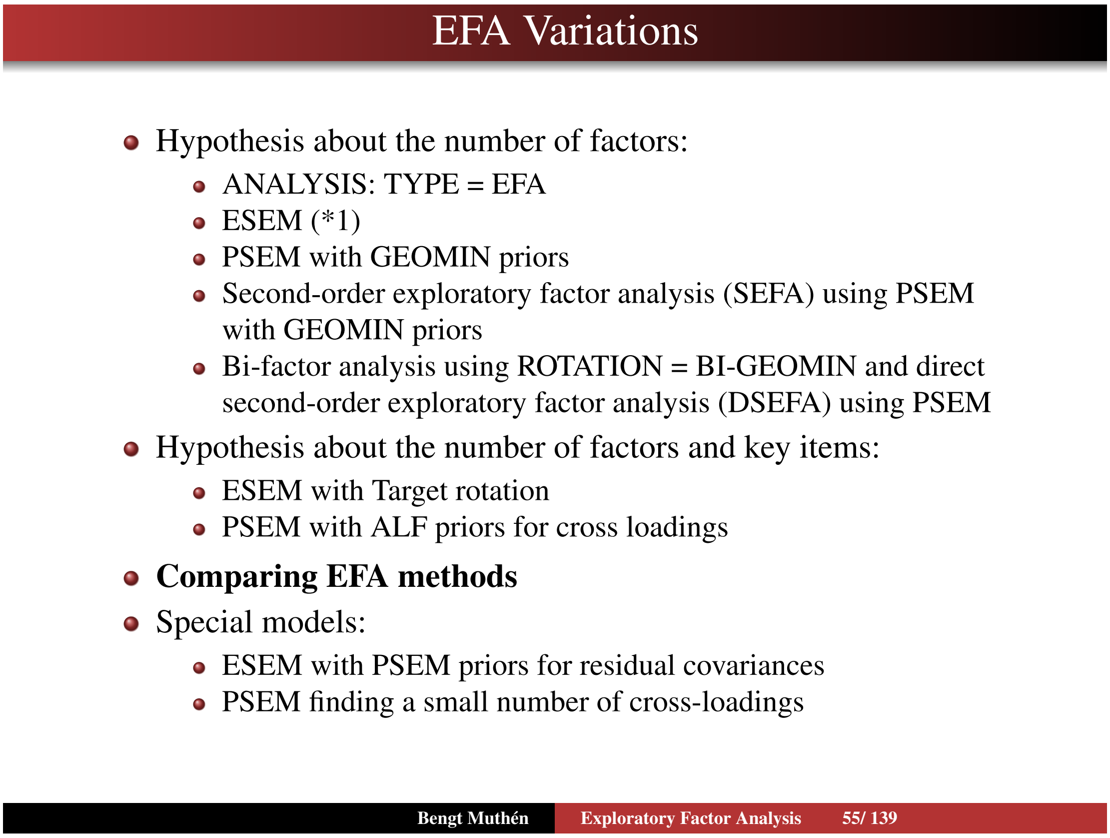
Bengt Muthén Exploratory Factor Analysis 54/ 139

Slide 54 shows the table of results. This is Table 4 of Section 4.4 of the 2024 PSEM paper. The critical loading λ_{61} is particularly sensitive to the method used. The target approach (ESEM Target) gets a large bias and poor coverage for this loading.

The overall conclusion is that PSEM ALF, PSEM GEOMIN, and ESEM GEOMIN perform best while PSEM Normal and ESEM Target perform the worst.

Problems with target rotation has been studied in the literature before, with suggestions to use an iterative target approach where target loadings that obtain large estimates are dropped as targets in a second target analysis. An overview of this topic is given in the Moore et al. (2015) paper given in the reference section.

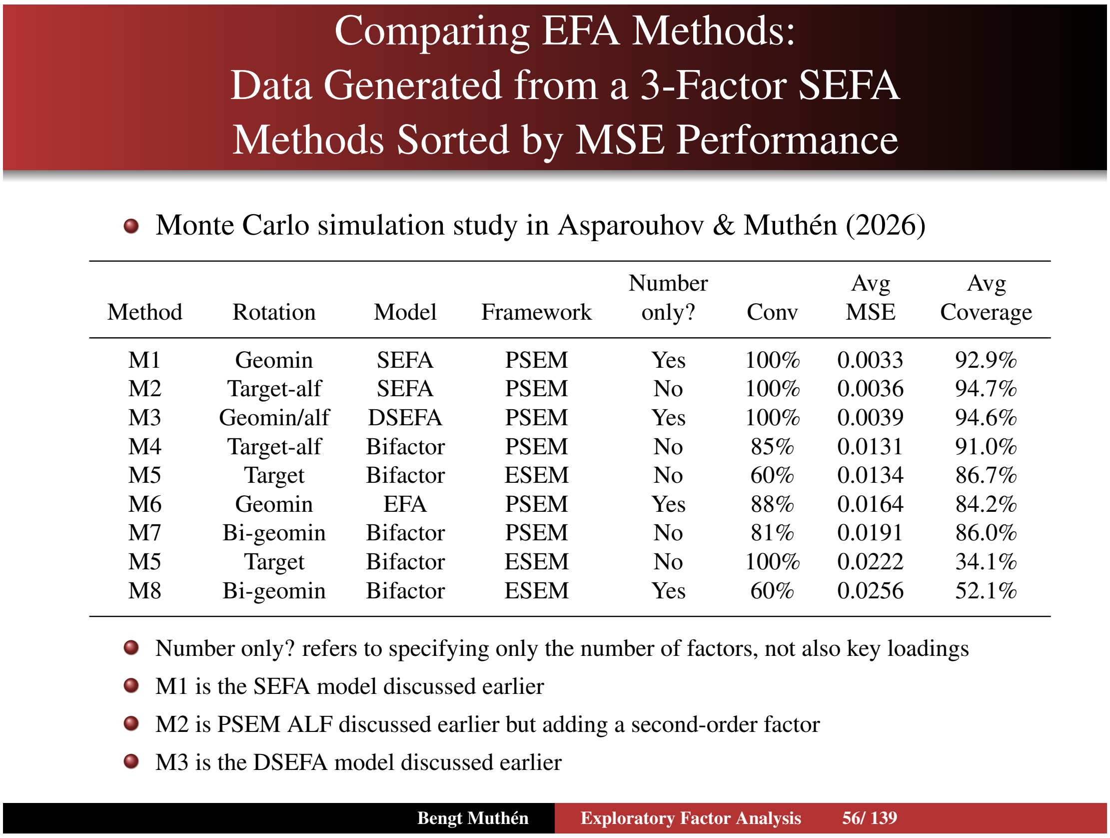
2.3 Comparing EFA methods



EFA Variations

- Hypothesis about the number of factors:
  - ANALYSIS: TYPE = EFA
  - ESEM (*1)
  - PSEM with GEOMIN priors
  - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
  - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- Hypothesis about the number of factors and key items:
  - ESEM with Target rotation
  - PSEM with ALF priors for cross loadings
- **Comparing EFA methods**
- Special models:
  - ESEM with PSEM priors for residual covariances
  - PSEM finding a small number of cross-loadings

Slide 55 returns to the EFA Variations overview. Here we turn to comparisons of exploratory second-order and bi-factor analyses based on simulation results. Target methods are also included.



Comparing EFA Methods: Data Generated from a 3-Factor SEFA Methods Sorted by MSE Performance

- Monte Carlo simulation study in Asparouhov & Muthén (2026)

Method	Rotation	Model	Framework	Number only?	Conv	Avg MSE	Avg Coverage
M1	Geomin	SEFA	PSEM	Yes	100%	0.0033	92.9%
M2	Target-alf	SEFA	PSEM	No	100%	0.0036	94.7%
M3	Geomin/alf	DSEFA	PSEM	Yes	100%	0.0039	94.6%
M4	Target-alf	Bifactor	PSEM	No	85%	0.0131	91.0%
M5	Target	Bifactor	ESEM	No	60%	0.0134	86.7%
M6	Geomin	EFA	PSEM	Yes	88%	0.0164	84.2%
M7	Bi-geomin	Bifactor	PSEM	No	81%	0.0191	86.0%
M5	Target	Bifactor	ESEM	No	100%	0.0222	34.1%
M8	Bi-geomin	Bifactor	ESEM	Yes	60%	0.0256	52.1%

- Number only? refers to specifying only the number of factors, not also key loadings
- M1 is the SEFA model discussed earlier
- M2 is PSEM ALF discussed earlier but adding a second-order factor
- M3 is the DSEFA model discussed earlier

Slides 56 and 57 show simulation results from the Asparouhov & Muthén 2026 paper: A unification of second-order and bi-factor EFA. The methods are listed in the order of performance from best to worst. The performance of the methods is judged by convergence percentage, average mean square error, and average coverage.

In the table on slide 56, the data are generated from a 3-factor SEFA and it is natural that this method, which is method M1, performs the best. Method M2 performs second best but needs not only specification of the number of factors but also locations of the targets.

Comparing EFA Methods:
Data Generated from a 4-Factor Bi-Factor Model
Methods Sorted by MSE Performance

- Monte Carlo simulation study in Asparouhov & Muthén (2026)

Method	Rotation	Model	Framework	Number only?	Conv	Avg MSE	Avg Coverage
M3	Geomin/alf	DSEFA	PSEM	Yes	100%	0.0069	93.8%
M4	Target-alf	Bifactor	PSEM	No	99%	0.0152	97.1%
M1	Geomin	SEFA	PSEM	Yes	98%	0.0167	93.4%
M5	Target	Bifactor	ESEM	No	97%	0.0255	87.2%
M8	Bi-geomin	Bifactor	ESEM	Yes	97%	0.0270	93.5%

- Number only? refers to specifying only the number of factors, not also key loadings
- M4 is the bi-factor PSEM ALF shown on the next slide
- M5 is the bi-factor target ESEM shown on slide 47

- M1 and M3 are among the top 3 in both tables and require only the number of factors
- M2 and M4 are among the top when key loadings are also specified

In the table of slide 57, the data are generated from a 4-factor bi-factor model so that method M3 performs best.

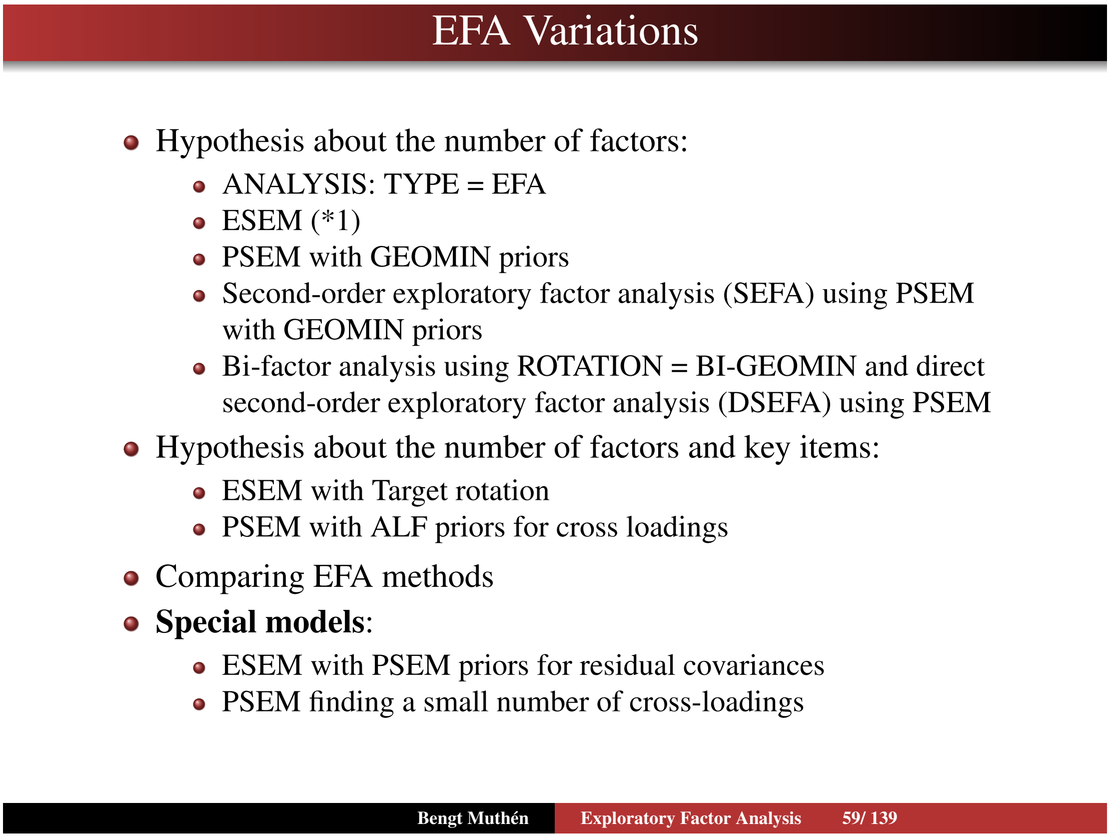
Drawing on the results from both tables, we conclude that the SEFA and DSEFA methods M1 and M3 are doing well - and importantly - require only the specification of the number of factors. Methods M2 and M4 do well when the location of key loadings are also specified. It is notable that M8 performs the worst in this simulation. It uses ROTATION = BI-GEOMIN which has to date been the most commonly used.

Input for Bi-Factor PSEM using ALF Cross-Loading Priors (M4)

```
ANALYSIS:      ESTIMATOR = MLR;
               ITERATIONS = 10000;
               CONVERGENCE = 0.000001;
               STARTS = 50;

MODEL:         ! 4 specific factors
               spatial BY
                 visual-flags*1
                 general-arithmet*0 (a1-a20);

               verbal BY
                 visual-flags*0 (b1-b4)
                 general-wordm*1
                 addition-arithmet*0 (c1-c15);

               speed BY
                 visual-wordm*0 (d1-d9)
                 addition-straight*1
                 wordr-arithmet*0 (e1-e11);

               memory BY
                 visual-straight*0 (f1-f13)
                 wordr-figurew*1
                 deduct-arithmet*0 (g1-g5);

               ! general factor:
               gen BY visual-arithmet*;

               spatial-gen WITH spatial-gen@0;
               spatial-gen@1;

MODEL PRIORS:  a1-g5 ~ ALF(0,1);

OUTPUT:        STANDARDIZED;
```

For completeness, slide 58 shows the input for bi-factor PSEM with ALF cross loadings priors. This is method M4 in the previous tables. The input refers to the 24-variable H&S data.

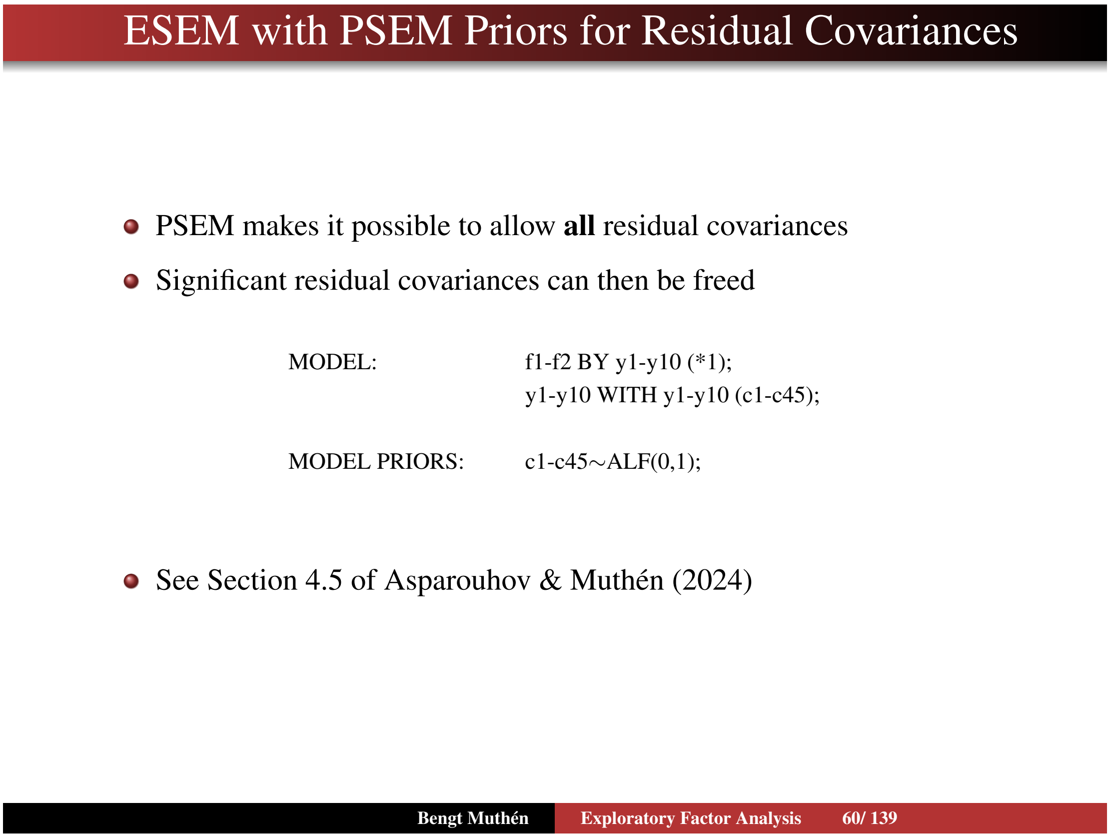
2.4 Special models



EFA Variations

- Hypothesis about the number of factors:
  - ANALYSIS: TYPE = EFA
  - ESEM (*1)
  - PSEM with GEOMIN priors
  - Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors
  - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) using PSEM
- Hypothesis about the number of factors and key items:
  - ESEM with Target rotation
  - PSEM with ALF priors for cross loadings
- Comparing EFA methods
- **Special models**:
  - ESEM with PSEM priors for residual covariances
  - PSEM finding a small number of cross-loadings

Slide 59 again shows the overview of EFA Variations. We now turn to the last 2 items on the list under Special models: ESEM where all residual covariances are included and PSEM searching for a small number of cross loadings.



ESEM with PSEM Priors for Residual Covariances

- PSEM makes it possible to allow **all** residual covariances
- Significant residual covariances can then be freed

```
MODEL:              f1-f2 BY y1-y10 (*1);
                    y1-y10 WITH y1-y10 (c1-c45);

MODEL PRIORS:       c1-c45∼ALF(0,1);
```

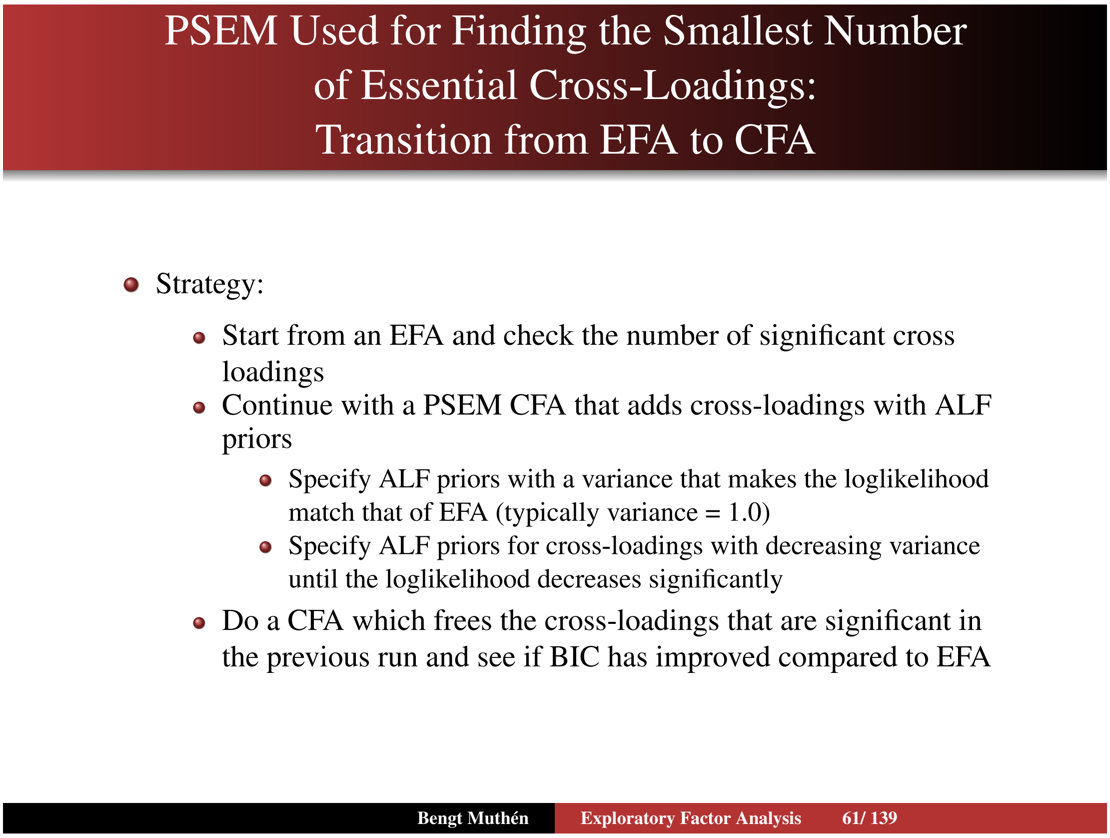
- See Section 4.5 of Asparouhov & Muthén (2024)

Slide 60 shows that it is possible to include residual covariances in an EFA by using PSEM. In the previous British Household Panel example we used ESEM to include some residual covariances. With PSEM, however, we can include all of them. This makes it possible to see which of them are significant and should be freed.

As the input shows, ALF priors can be applied to all the residual covariances. These priors push the residual covariances toward the ALF prior mean of zero. If the dataset clearly requires a residual covariance, it will overwhelm the prior and show a substantial non-zero residual covariance estimate.

The technical background for this is discussed in section 4.5 of the 2024 Asparouhov-Muthén article on PSEM.

It is not recommended to use this approach in conjunction with more complex models such as SEFA and DSEFA. The alternative approach of using modindices with regular EFA was presented on slide 42.



PSEM Used for Finding the Smallest Number of Essential Cross-Loadings: Transition from EFA to CFA

- Strategy:
  - Start from an EFA and check the number of significant cross loadings
  - Continue with a PSEM CFA that adds cross-loadings with ALF priors
    - Specify ALF priors with a variance that makes the loglikelihood match that of EFA (typically variance = 1.0)
    - Specify ALF priors for cross-loadings with decreasing variance until the loglikelihood decreases significantly
  - Do a CFA which frees the cross-loadings that are significant in the previous run and see if BIC has improved compared to EFA

Slide 61 shows a different use of PSEM. It can answer the question of which is the smallest set of essential cross loadings when transitioning from an EFA to a CFA. BIC is used as the basis for the decision.

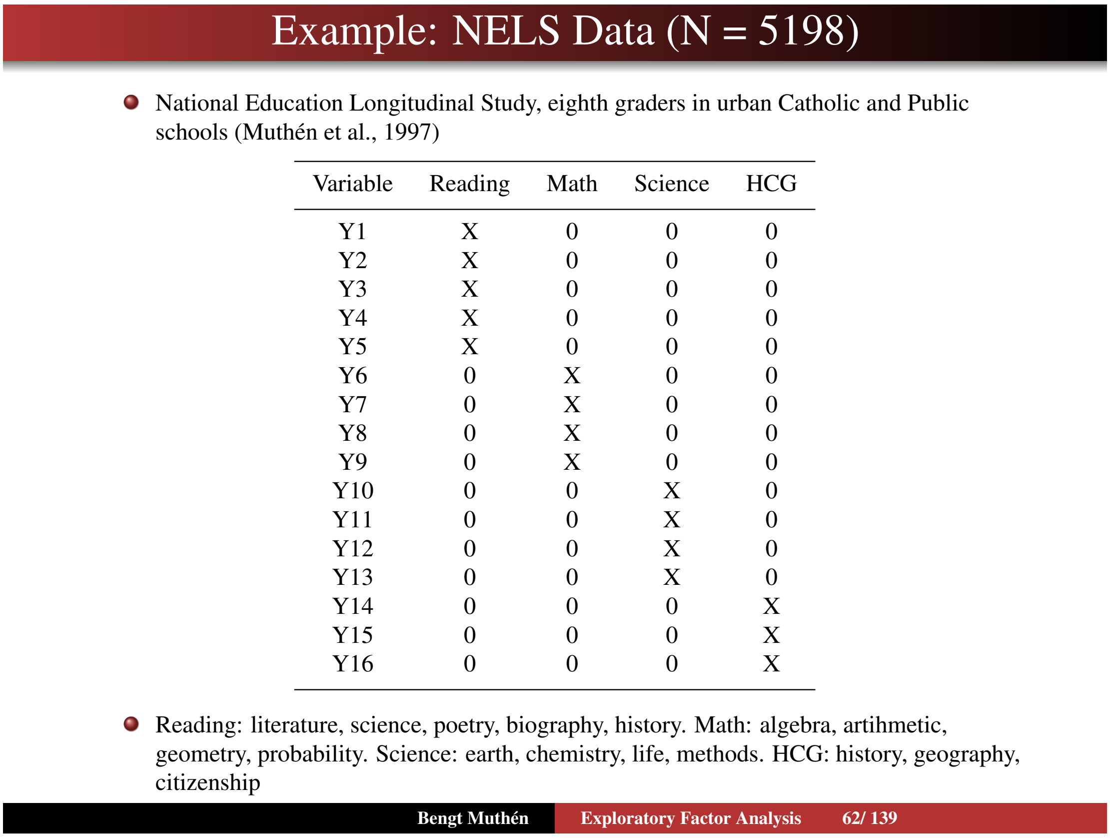
The strategy is to start from an EFA where you check the number of significant cross loadings.

You then continue with a PSEM CFA that adds cross-loadings with ALF priors.

- First you specify ALF priors with a variance that makes the loglikelihood match that of EFA (typically variance = 1.0).

- Then you specify these ALF priors for cross-loadings with decreasing variance until the loglikelihood decreases substantially.

Finally, you do a CFA where you free the cross-loadings that are significant in the previous run and see if BIC has improved compared to EFA.



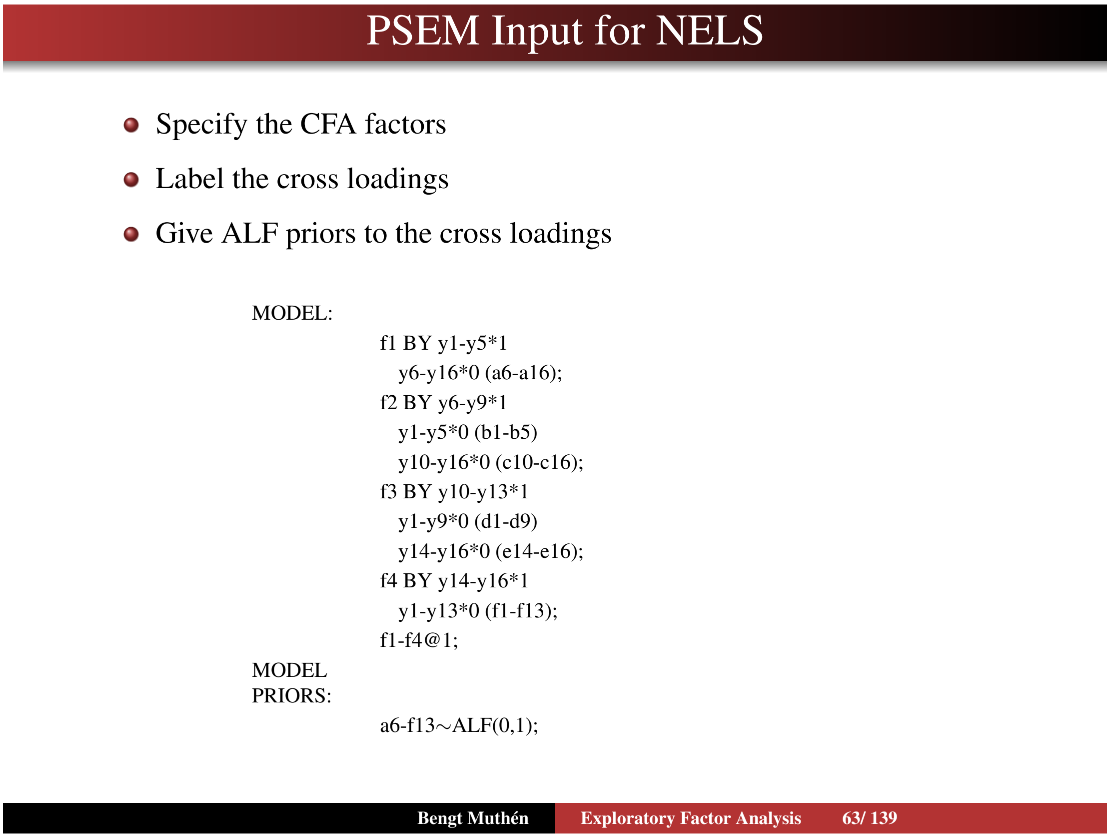
Example: NELS Data (N = 5198)

- National Education Longitudinal Study, eighth graders in urban Catholic and Public schools (Muthén et al., 1997)

Variable	Reading	Math	Science	HCG
Y1	X	0	0	0
Y2	X	0	0	0
Y3	X	0	0	0
Y4	X	0	0	0
Y5	X	0	0	0
Y6	0	X	0	0
Y7	0	X	0	0
Y8	0	X	0	0
Y9	0	X	0	0
Y10	0	0	X	0
Y11	0	0	X	0
Y12	0	0	X	0
Y13	0	0	X	0
Y14	0	0	0	X
Y15	0	0	0	X
Y16	0	0	0	X

- Reading: literature, science, poetry, biography, history. Math: algebra, artihmetic, geometry, probability. Science: earth, chemistry, life, methods. HCG: history, geography, citizenship

Slide 62 shows the expected loading pattern for a measurement instrument used in the National Education Longitudinal Study of eighth graders in the US. The sample size is 5,198. Four factors are expected. The 16 factor indicators are described at the bottom of the table.

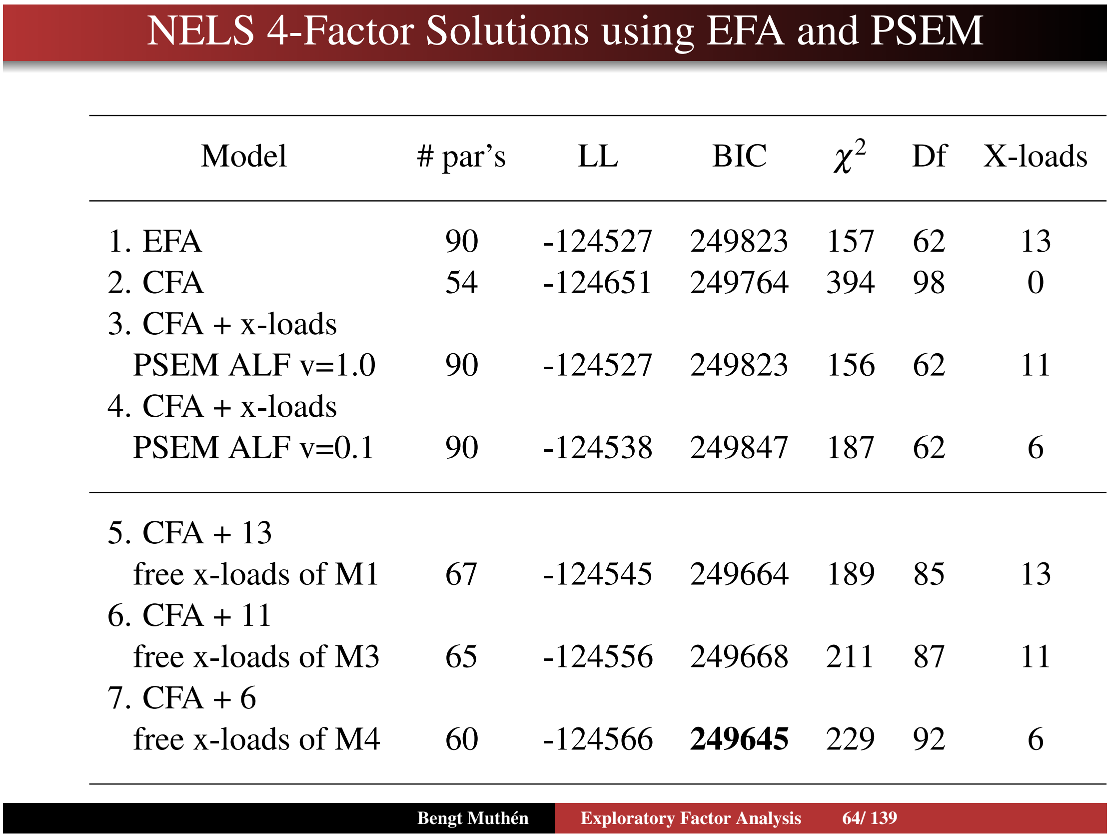


PSEM Input for NELS

- Specify the CFA factors
- Label the cross loadings
- Give ALF priors to the cross loadings

```
MODEL:
              f1 BY y1-y5*1
                 y6-y16*0 (a6-a16);
              f2 BY y6-y9*1
                 y1-y5*0 (b1-b5)
                 y10-y16*0 (c10-c16);
              f3 BY y10-y13*1
                 y1-y9*0 (d1-d9)
                 y14-y16*0 (e14-e16);
              f4 BY y14-y16*1
                 y1-y13*0 (f1-f13);
              f1-f4@1;
MODEL
PRIORS:
              a6-f13~ALF(0,1);
```

Slide 63 shows the input for the PSEM part of the exploration of essential cross loadings.
You specify the CFA factor pattern, label the cross loadings, and give ALF priors to those cross loadings.



NELS 4-Factor Solutions using EFA and PSEM

Model	# par's	LL	BIC	χ^2	Df	X-loads
1. EFA	90	-124527	249823	157	62	13
2. CFA	54	-124651	249764	394	98	0
3. CFA + x-loads PSEM ALF v=1.0	90	-124527	249823	156	62	11
4. CFA + x-loads PSEM ALF v=0.1	90	-124538	249847	187	62	6
5. CFA + 13 free x-loads of M1	67	-124545	249664	189	85	13
6. CFA + 11 free x-loads of M3	65	-124556	249668	211	87	11
7. CFA + 6 free x-loads of M4	60	-124566	**249645**	229	92	6

Slide 64 shows a series of models using EFA, CFA, and CFA with PSEM. Of particular importance are the loglikelihood (LL) and BIC values.

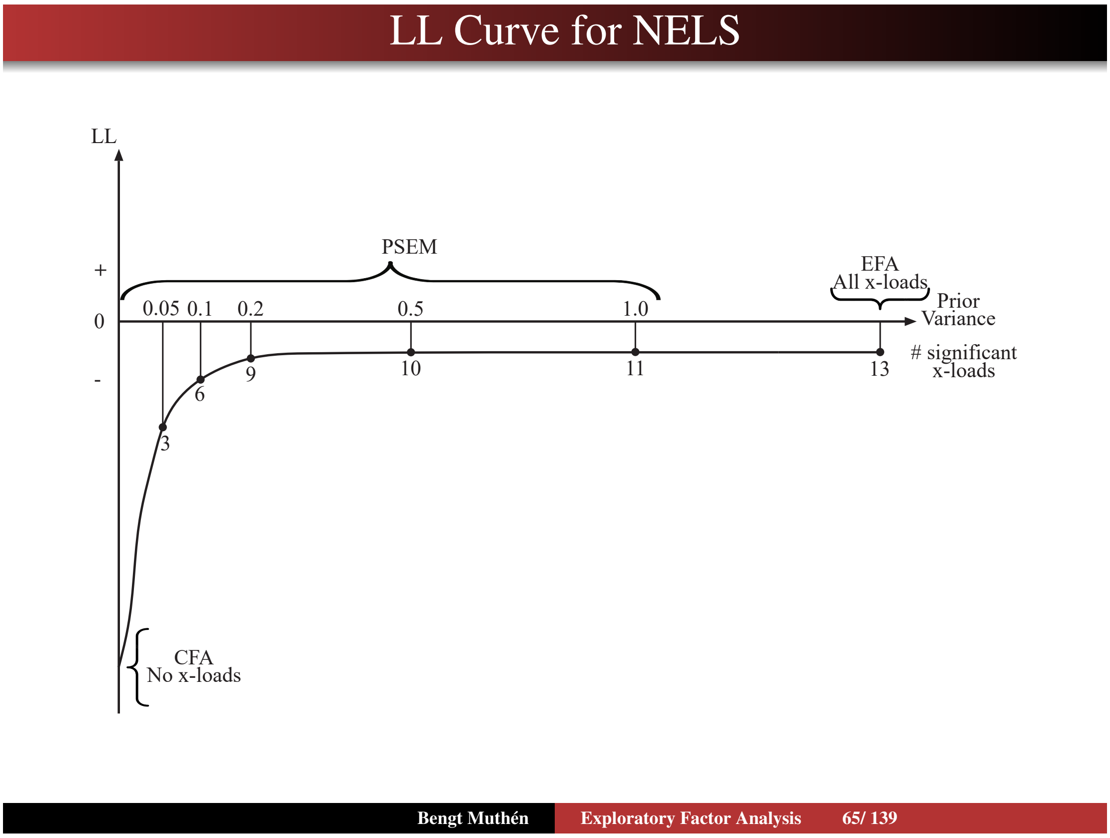
Model 1 is a 4-factor EFA which has 13 significant cross loadings.

Model 2 is a regular CFA. The CFA model has better (lower) BIC than the EFA.

Model 3 is a CFA adding all cross-loadings using ALF priors. The variance 1 results in the same model fit as the Model 1 EFA as evidenced by the log likelihood (LL). This results in 11 significant cross loadings, representing a modest reduction from the 13 found with EFA. Having found this model fit equivalence with EFA, smaller variances for the ALF priors can be explored.

Model 4 uses ALF prior variance of 0.1. As expected, this worsens the log likelihood. To gauge the magnitude of the drop, it is useful to compare it to the distance between the EFA and CFA log likelihoods. In those terms, the drop is only 8.9%. For this model, the number of significant cross loadings is only 6. Models with small log likelihood drops and a number of significant cross loadings smaller than that of EFA provide a basis for a follow-up CFA that may have a competitive BIC as illustrated in the bottom part of the table.

The bottom part of the table adds significant cross loadings to the regular CFA Model 2 based on models 1, 3, and 4. These 3 models give better (lower) BIC values than any of the models in the top part. Of particular importance is that Model 7 with only 6 cross loadings gives the best BIC and is therefore the model of choice.



LL Curve for NELS

Slide 65 shows a log likelihood curve for the different analyses on the previous slide. At top right is the EFA model which has all cross loadings of which 13 are significant. It can be seen as a PSEM model where the ALF prior variances are very large so that the priors essentially have no effect.

Bottom left is the CFA model which has no cross loadings. In between, there is a series of PSEM models with varying ALF prior variance and corresponding number of significant cross loadings. The plot shows the chosen model which has prior variance 0.1 with 6 significant cross loadings, representing the 8.9% drop in the log likelihood, and receiving the best BIC.