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

EFA Variations

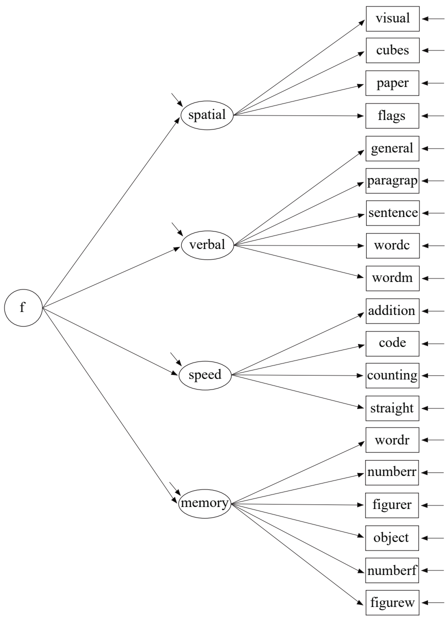
- Hypothesis about the number of factors:
  - ANALYSIS: TYPE = EFA
  - ESEM (*1)
  - PSEM with GEOMIN priors
  - **Second-order exploratory factor analysis (SEFA) using PSEM with GEOMIN priors**
  - Bi-factor analysis using ROTATION = BI-GEOMIN and direct second-order exploratory factor analysis (DSEFA) 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

Bengt Muthén | Exploratory Factor Analysis | 21/ 139

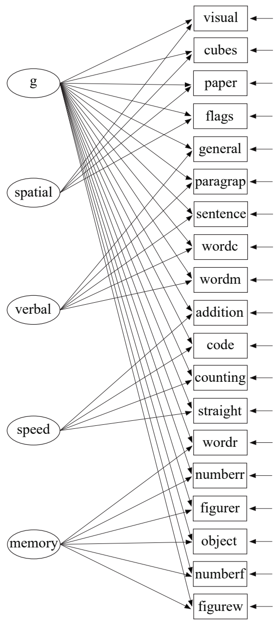
Slide 21 returns to the overview of EFA Variations. We now turn to the topic of second-order factor analysis where we will also introduce the SEFA model using PSEM with GEOMIN priors.

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);
```

Bengt Muthén | Exploratory Factor Analysis | 26/ 139

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.