A note on second-order EFA

Tihomir Asparouhov and Bengt Muthén

Mplus

January 15, 2026

Abstract

As experience with hierarchical exploratory factor analysis has accumulated, it has become necessary to revisit, update, and correct several aspects of this modeling approach, including its estimation, interpretation, and the expectations of what the model can accomplish. With these refinements, the model can assume its rightful place as a serious alternative to bi-factor EFA models and to EFA models with high factor correlations.

1 Introduction

Hierarchical Exploratory Factor Analysis (HEFA) was discussed in Asparouhov and Muthén (2024, 2025) in the context of the Penalized Structural Equation Modeling (PSEM) framework. The HEFA model can also be viewed as a second-order EFA model where the first-order factor analysis is an EFA/CFA model and the second-order factor model is also an EFA/CFA model where the factors from the first-order analysis are used as indicators in the second-order analysis. The model is given by the following equations:

$$Y = \nu + \Lambda_1 F + \varepsilon, \tag{1}$$

$$F = \Lambda_2 \eta + \xi, \tag{2}$$

where Λ_1 and Λ_2 can be unconstrained matrices of loadings as in EFA or be structured as in CFA, F is a vector of first-order factors and η is a vector of second order factors.

This note is intended to provide corrections and clarifications to Section 6.2 of Asparouhov and Muthén (2024), which discusses the estimation of HEFA models where the first-order analysis is exploratory. Additional practical experience with this HEFA model has developed into an improved PSEM based estimation, better interpretation of the model, and how it compares to other models. Below we discuss all of these aspects in detail. The most common model to estimate among these HEFA models is the case where there are several exploratory first-order factors and one second-order factor. We will primarily discuss this simple case here; however, all conclusions apply to second-order models with more than one factor, models where the second-order analysis is also exploratory, and models which also include observed second-order indicators. This paper also serves as another illustration of the power of the PSEM framework in Mplus.

In Morin et al. (2016), the HEFA model is discussed with a two-stage estimation procedure referred to as the EWC approach, formulated within the ESEM framework of Asparouhov and Muthén (2009). Ironically, that estimation method also required revision for a reason very similar to the PSEM-based estimation we revise here (see Morin and Asparouhov, 2018). We currently consider the PSEM-based approach to be superior to the EWC approach because it is based on a single step rather than two steps, is less time-consuming to set up, and is less prone to estimation errors due to its simpler setup. Therefore, the updates and corrections we provide here are essential and should result in better utilization of the HEFA model.

2 Estimation

The HEFA model is estimated with PSEM where geomin(Λ_1) is set as the penalty function. Here geomin is the rotation function associated with the geomin rotation.

Figure 13 in the supplemental materials of Asparouhov and Muthén (2024) illustrates how the model can be estimated. Simulation studies are included in Asparouhov and Muthén (2024), which also show that model estimation works well. However, some practical experience with this model has revealed a weakness in the estimation: the estimation often fails to converge, particularly for smaller sample size situations. For example, the Figure 13 model has a convergence rate of 100% with sample size $N = 2000$, but the convergence rate with $N = 500$ is 75%, and with $N = 300$ it is 46%. In this note we propose a new estimation method which has 100% convergence at $N = 2000$, 500, and 300. The key change in the estimation is how the first-order factor scales are set. In Figure 13, the residual variance is fixed to 1, i.e., $\text{Var}(\xi) = 1$. What we propose as a solution to the convergence issues is to replace that with $\text{Var}(F) = 1$, i.e., the factor scale is set by fixing the total variance of F to 1 as in single-order EFA.

If we use only the residual variance to set the scale, the penalized likelihood has multiple local maxima that compete with the true values. Geomin optimization occurs when loadings are near zero. The alternate local maxima occurs when, for a particular first-order factor, the first-order loadings are divided by a large number while the second-order loadings are multiplied by that large number. In this process, the overall covariance structure remains the same except that in the second-order factor model, the second-order factor is identical to that first-order factor. The data log-likelihood decrease counterbalances the gain from the penalty optimization. These local maxima may also be global maxima in some situations, but even if they are not, it is quite easy for the optimization process to diverge towards these alternate solutions. If there are m first-order factors, essentially there are m alternate local maximas that the optimization can diverge towards. These problems are easily resolved when we use the total variance to fix the scales in the factor models. Such a setup separates the two sets of loadings and no multiplicative constant can pass through.

Because the total first-order factor variances are not model parameters, to fix these to 1 we must use model constraints in Mplus. Assuming one second-order factor with a variance fixed to 1, the total variance of the first-order factor is

$$\mathrm{Var}(F_i) = \lambda_{2i}^2 + \mathrm{Var}(\xi_i). \quad (3)$$

The most efficient way to constrain this total variance to 1 in Mplus is to constrain the residual variance as

$$\mathrm{Var}(\xi_i) = 1 - \lambda_{2i}^2. \quad (4)$$

Figure 1 illustrates this setup for a simulation study with four first order factors and one second order factor. Note here that the first-order factor residual variances' true values are chosen so that the above constraint is satisfied. As in Figure 13 of Asparouhov and Muthén (2024), PSEM Geomin priors are used for the rotation of Λ_1. The results of this simulation study for a portion of the first order loadings and all of the second-order loadings are given in Figure 2. The results indicate that the parameters are recovered reasonably well with a sample size of $N = 500$. As discussed in the next section, the chi-square test of fit for the second-order EFA with m first-order factors is the same as the chi-square test of fit for the standard EFA with m factors. In this example, the average chi-square value is 117, and with 116 degrees of freedom this yields a rejection rate of 3%.

Figure 3 shows the input file for a single data set that can be used with real data where starting values are not available. The key feature that is needed now is the option `STARTS=50`. Starting values for the first-order loadings can be obtained from the standard EFA; however, even without starting values the algorithm will converge given enough randomly generated starting values. When using random starting values, neither the sign of the loadings nor the order of the factors is presented in an organized fashion as it is done in the Mplus EFA estimator. That means that some factors may appear with all negative main loadings, and the factors themselves may appear in random order. Thus, if starting values are available or a structure is hypothesized, the PSEM estimator should be used with starting values. This is illustrated in the empirical example section.

It is argued in Asparouhov and Muthén (2024) that the second-order EFA is more difficult to estimate than other EFA models or bi-factor EFA models. It turns out that this was an artifact of the suboptimal estimation method used at the time and it is no longer true when we use the total first-order factor variance to set the scale. With the new estimation method, the second-order EFA is about as difficult to estimate as any other EFA model. This appears to be true not just for simulated data but also in real data analysis.

Figure 1: Simulation study for a second order EFA

```
MONTECARLO:
    NAMES = y1-y20;
    NOBSERVATIONS = 500;
    NREPS = 100;

MODEL POPULATION:
    f1 BY y1*0.7 y2*1.3 y3-y4*0.8 y5*0.3;
    f2 BY y5*0.6 y6*0.7 y7*0.5 y8*0.8 y9*1 y10*0.5 ;
    f3 BY y11*0.7 y12*1 y13-y15*0.4;
    f4 BY y15*0.6 y16*0.7 y17*1 y18*0.5 y19*1 y20*0.5;
    f1*0.84 f2*0.75 f3*0.75 f4*.64;
    y1-y20*1;
    f0 by f1*0.4 f2*0.5 f3*0.5 f4*0.6; f0@1;

MODEL:
    f1 BY y1*0.7 y2*1.3 y3-y4*0.8 y5*0.3 y6-y20*0(a1-a20);
    f2 BY y1-y4*0 y5*0.6 y6*0.7 y7*0.5 y8*0.8 y9*1 y10*0.5 y11-y20*0 (a21-a40);
    f3 BY y1-y10*0 y11*0.7 y12*1 y13-y14*0.4 y15*0.4 y16-y20*0 (a41-a60);
    f4 BY y1-y14*0 y15*0.6 y16*0.7 y17*1 y18*0.5 y19*1 y20*0.5 (a61-a80);
    f1*0.84 f2*0.75 f3*0.75 f4*.64 (v1-v4);
    y1-y20*1;
    f0 by f1*0.4 f2*0.5 f3*0.5 f4*0.6 (l1-l4); f0@1;

model constraint:
v1=1-l1*l1;
v2=1-l2*l2;
v3=1-l3*l3;
v4=1-l4*l4;

MODEL PRIORS: a1-a80~Geomin(4,0.1);

analysis: iter=10000; conv=0.000001;
```

Figure 2: Results of simulation study for a second-order EFA

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
F1 BY							
Y1	0.700	0.6878	0.0640	0.0609	0.0042	0.920	1.000
Y2	1.300	1.2919	0.0760	0.0786	0.0058	0.950	1.000
Y3	0.800	0.7895	0.0547	0.0634	0.0031	0.960	1.000
Y4	0.800	0.7966	0.0643	0.0634	0.0041	0.960	1.000
Y5	0.300	0.3076	0.0581	0.0619	0.0034	0.960	1.000
Y6	0.000	0.0136	0.0458	0.0527	0.0023	0.980	0.020
Y7	0.000	0.0047	0.0467	0.0542	0.0022	0.960	0.040
Y8	0.000	0.0027	0.0464	0.0522	0.0021	1.000	0.000
Y9	0.000	0.0038	0.0424	0.0501	0.0018	1.000	0.000
Y10	0.000	0.0049	0.0494	0.0535	0.0024	0.980	0.020
Y11	0.000	-0.0017	0.0485	0.0530	0.0023	0.990	0.010
Y12	0.000	0.0042	0.0567	0.0517	0.0032	0.970	0.030
Y13	0.000	0.0042	0.0522	0.0544	0.0027	0.960	0.040
Y14	0.000	0.0133	0.0545	0.0543	0.0031	0.960	0.040
Y15	0.000	0.0032	0.0466	0.0449	0.0022	0.990	0.010
Y16	0.000	0.0097	0.0525	0.0521	0.0028	0.970	0.030
Y17	0.000	0.0080	0.0451	0.0482	0.0021	0.980	0.020
Y18	0.000	0.0001	0.0505	0.0528	0.0025	0.970	0.030
Y19	0.000	0.0057	0.0469	0.0485	0.0022	0.980	0.020
Y20	0.000	0.0095	0.0557	0.0531	0.0032	0.970	0.030
F0 BY							
F1	0.400	0.3756	0.0715	0.0813	0.0057	0.970	0.990
F2	0.500	0.4674	0.0739	0.0893	0.0065	0.980	0.990
F3	0.500	0.4688	0.0851	0.0950	0.0082	0.950	0.990
F4	0.600	0.5703	0.0763	0.0976	0.0066	0.970	0.990

Figure 3: Second order EFA input file for real data analysis

```
variable: NAMES = y1-y20;
data: file=1.dat;

MODEL:
    f1-f4 BY y1-y20*(a1-a80);
    f1-f4 (v1-v4);
    f0 by f1-f4* (l1-l4); f0@1;

model constraint:
v1=1-l1*l1;
v2=1-l2*l2;
v3=1-l3*l3;
v4=1-l4*l4;

MODEL PRIORS: a1-a80~Geomin(4,0.1);

analysis: iter=10000; conv=0.000001; starts = 50;
```

3 Interpretation

A second-order EFA with m_1 first-order factors and m_2 second-order factors has the same fit as the EFA with m_1 factors, regardless of whether the second-order model is an EFA or a CFA model. The second order EFA model is equivalent to a factor analysis model with a factor correlation matrix Ψ implied by the second order factor analysis model. It is well known that EFA with m_1 factors yields the best data fit among all factor models with m_1 factors. Thus the second order EFA with m_1 first order factors in terms of log-likelihood can not exceed the log-likelihood of the EFA with m_1 factors. Now we will show that the opposite is true: the maximum likelihood of the second order EFA is at least as high as that for the EFA model. Every correlation matrix Ψ can be represented as $\Psi = HH^T$ where H is an oblique rotation matrix. If Λ_0 is the loading matrix of the unrotated EFA solution with unit matrix as the factor correlation matrix, we can rotate the unrotated model using H, to obtain a factor model, with factor correlation matrix $\Psi = HH^T$ and loading matrix $\Lambda_0 H^{-1}$. Since the loading matrix in the second order EFA is unconstrained, this solution is among those that are considered in the second order EFA log-likelihood optimization. To rephrase, for every set of parameters in the second order model, there is a loading matrix of the first order factor model, which yields the same optimal log-likelihood as the unrotated EFA model with m_1 factors. We conclude that in terms of data fit, a second-order EFA with m_1 first-order factors is equivalent to the EFA model with m_1 factors.

Every PSEM model estimation is based on a pair of models in addition to the penalty function. The first model is the actual PSEM model we are trying to estimate. This model is typically unidentified if estimated without the penalty. This means that a multidimensional space of model parameters yields the same optimal data fit. In that space, we identify the PSEM model estimates by minimizing the penalty. Usually in that space we can identify a well-known model that can be used as a reference. The model is identified by other means not related to the penalty, such as fixing some of the parameters. This model is referred to as the null model and its primary function is to ensure that the PSEM model yields the same log-likelihood, while varying the penalty weight, see Asparouhov and Muthén (2024) for details. For the second-order EFA, the null model is the unrotated EFA model. However, since the unrotated EFA model and the rotated EFA model have the same log-likelihood, we can regard the EFA model with m_1 factors as the null model for the second-order EFA with m_1 first order factors and m_2 second order factors. This means that the geomin penalty weight is reduced to the point where the penalized likelihood optimization yields a data fit comparable with the EFA model. In the example of Figure 1, a comparable fit is obtained with geomin prior variance of 0.1. It is interesting to note here that m_2 does not affect the data fit. The role of m_2 is only regarding the allowed rotations. The rotation in the second-order EFA is neither oblique nor orthogonal. The rotations for the second-order EFA are a subset of the oblique rotations which conform with the second-order model. This also means that we cannot evaluate the second-order model using the data fit, just like we cannot pick between oblique or orthogonal EFA based on the data fit: they have identical data fit.

The advantage of the oblique EFA over the orthogonal EFA is that a simpler loading structure can be obtained with the oblique EFA. The drawback of it is that the factor correlation matrix is unrestricted and can be viewed as more complex. This perspective allows us to place the value of the second-order EFA in that framework. A second-order EFA will provide a simpler loading structure than an orthogonal EFA, and it gives us an opportunity to model the first-order factor correlation structure with a simpler/more parsimonious and more interpretable second-order factor model. In addition, the second-order factor can also provide a valuable interpretation for the entire model as an inherent personal characteristic affecting the first-order factors. Furthermore, using a second-order factor provides for a parsimonious model when relating to covariates or when used as a predictor of other variables.

A second-order EFA lies between the orthogonal EFA and the oblique EFA on the scale of how simple the loading structure is for the indicators. It also lies in between the two traditional EFA models in terms of simplicity of the factor correlation. The main question that needs to be answered is when a second-order EFA should be preferred to the traditional orthogonal and oblique EFA. If the second-order EFA model's first-order loading structure Λ_1 is better and more interpretable than the orthogonal EFA loading structure and is comparable (not substantially worse) than the oblique model, it should be preferred and be considered the best among the three models.

Consider the example in Figure 3. The data is generated using the simulation study in Figure 1. There are 22 first-order loadings that are not zero. The orthogonal model yields a factor structure with 28 significant loadings. On the other hand, the second-order EFA and the oblique EFA yield nearly identical loading structure: a total of 24 significant loadings (2 spurious cross-loadings are found with both models). We conclude that restricting the oblique rotations to a second-order factor model does not result in worse loading structure and therefore this is evidence for the usefulness of the second-order factor which explains the correlation between the first-order EFA factors. Thus, in this example, the second-order EFA model should be preferred.

In Asparouhov and Muthén (2024), it is argued that when (1) and (2) are combined, the equation

$$Y = \nu + \Lambda_1 \xi + \Lambda_1 \Lambda_2 \eta + \varepsilon \tag{5}$$

resembles a bi-factor model. To some extent this is true because if the second-order EFA has m_1 first-order factors and m_2 second-order factors, we end up with a factor model with $m_1 + m_2$ factors: ξ and η, which are independent, and the ξ factors are also independent among each other. Note that the factors ξ are the factor residuals of the first order factors F. The above model is indeed a bi-factor model with η being the general factors and ξ the specific factors. However, this is not the bi-factor EFA model with m_1 specific and m_2 general factors. This is because the loading matrices are highly structured and not unrestricted like in EFA. Most importantly, however, if we are looking for a model fit equivalence, this second-order EFA is equivalent to a bi-factor EFA with a total of m_1 factors: $m_1 - 1$ specific and 1 general, or $m_1 - 2$ specific and 2 general, etc. This equivalence has nothing to do with the above equation at all. It is simply a conclusion that we can make from the fact that the second-order EFA is equivalent in terms of data fit to any EFA model with m_1 factors, including the bi-factor model with a total of m_1 factors.

In summary, second order EFA, just like the orthogonal EFA, compensates for its restrictive factor correlation structure with a few more cross-loadings (if needed) as compared to the oblique EFA.

In essence, we have three competing models: oblique EFA, second-order EFA, and bi-factor EFA. A natural question arises: is the precision of the loading estimates (i.e., MSE) affected by or does it benefit from a more restrictive factor structure? The answer appears to be no, or at least any effect is not of substantive importance. For example, the simulation study presented in Figure 1 yields identical loadings within each replication (and therefore identical MSE) for geomin second-order EFA and geomin oblique EFA. This result is expected because the true data-generating model is a second-order EFA. In general, however, this will not be the case, and oblique EFA is expected to produce a simpler—or at minimum, no more complex—loading matrix than second-order EFA. The converse is also true: if the loadings between the two models do not change substantially, we can take this as evidence that the second-order EFA represents the true model.

Oblique EFA may sometimes produce factor correlations that are too high to reliably distinguish between factors. Second-order EFA can address this problem by modeling the high factor correlations as a general feature and obtaining independent and distinguishable factor residuals ξ. EFA with high factor correlations is associated with multiple rotated solutions. That is, multiple local minimums of the geomin optimization can be found in close proximity. The second-order EFA, which has a more restrictive domain of rotations to optimize over, will certainly lead to fewer such local minimum situations. When sample size is small or moderate, geomin local minimum problems may manifest as lack of replicability. One small sample EFA result may differ dramatically from another small sample EFA result simply because small changes in the sample variance covariance may lead to an alternative geomin minimum. Thus, we expect that another advantage of the second-order EFA is the stability of the rotation.

To compare the bi-factor EFA model to the oblique EFA and second-order EFA models, we conduct the Figure 1 simulation study with the bi-geomin rotation using three specific factors and one general factor (which gives the correct number of factors). The results are somewhat awkward. The first three factors are recovered as specific, but the fourth factor is estimated as general, with all indicators loading on it. The model is awkward because the last five indicators have no specific factor. Typically, in a bi-factor model, each indicator loads on both a specific factor and a general factor. Thus, if a bi-factor model leaves many indicators without specific factors, this can be viewed as evidence that second-order EFA should be explored instead.

In practice, it is not unusual for bi-factor models to be presented where certain indicators are considered ”general ability” indicators, i.e., indicators without specific factors. For example, Harman (1967) presents a bi-factor model for the Holzinger and Swineford (1939) data where the last 5 indicators are considered general factor indicators, see Table 7.6 on page 129. Clearly this topic is somewhat subjective. One point of view is that general ability indicators would be hard to define in an unbiased way, and some may find it quite objectionable if for example items designed as "math ability" factor indicators, within a bi-factor EFA analysis, end up as "general ability" indicators. This would skew the "general ability" feature in favor of mathematics. From that perspective, the HEFA model offers an alternative interpretation that some might find less biased and less controversial. In addition, the HEFA model structure accommodates "general ability" indicators more naturally. If such indicators are truly designed as "general ability" indicators, these belong in equation (2) and not in equation (1). Such indicators should be used directly as indicators for the general factor rather than in the rotation measuring domain specific factors.

For the Figure 1 simulation study, a bi-factor model with five factors does provide a general factor and four correctly loaded specific factors, but such a model has an incorrect number of factors. This suggests that an additional advantage of second-order EFA is its ability to correctly parse a general underlying feature with fewer factors than a bi-factor model—that is, it provides a simpler model.

The bi-factor method has a long history, dating back to Holzinger and Swineford (1937), who promoted it as a confirmatory factor analysis (CFA) model that was straightforward to compute manually. More recently, the method has gained popularity because of its ability to separate a general factor from specific factors (Reise, 2012). In exploratory factor analysis (EFA) settings, bifactor estimation has been available since the work of Jennrich and Bentler (2011, 2012). However, these advantages of the bi-factor method over the second-order approach are now largely eliminated. Second-order EFA is more naturally embedded within the structural equation modeling (SEM) framework and may therefore be regarded as the superior method for extracting a general factor in EFA. A literature review summary and valuable discussion on the comparison of bi-factor and second-order modeling strategies is available in Morin et al. (2016). A variety of different points of view have been presented. Here, we want to reiterate that in EFA settings, the equivalence between these models holds only under specific and unexpected conditions. A second-order EFA with m first-order factors and one second-order factor is equivalent to the bi-factor EFA model with $m-1$ specific factors and one general factor. It is not the equivalent to the bi-factor model with m specific factors and one general factor, despite what simple factor counting or equation substitution might misleadingly suggest.

4 An empirical example

We illustrate the second-order EFA method with data from Holzinger and Swineford (1939). Test scores on 26 different mental ability measures were obtained from a total of 301 7th- and 8th-grade students in two schools. As in Muthén and Asparouhov (2012), we use 19 of the items that are hypothesized to measure four domains: spatial abilities, verbal abilities, speed, and memory. Holzinger and Swineford (1939) hypothesized the existence of a general factor affecting all test scores, and thus second-order EFA is a suitable modeling approach.

Figure 4 shows the input file for estimating the model. A larger number of iterations, stricter convergence criteria, and a larger number of random starting values specified in the `ANALYSIS` command are essential. Alternatively, we can bypass these requirements by using EFA starting values—that is, the loadings from the EFA geomin rotation as starting values for the second-order EFA. Figure 5 shows how to obtain such starting values. The figure contains only the commands that differ from Figure 4. This is an EFA model estimation using the ESEM framework in Mplus. The `(*1)` label in the measurement model is used to specify EFA. The `SVALUES` output option yields an Mplus model command where the model estimates from the EFA are placed as starting values. The loading parameter labels (`a1`–`a76`) in Figure 5 are essential and will also be included in the `SVALUES` output. We need not only the starting values but also the parameter labels, as these will be used to specify the rotation criteria in `MODEL PRIORS` for the second-order EFA. Entering the parameter labels and starting values manually would be time-consuming, as there are 76 loading parameters. Note that we will use only the loadings portion of the `SVALUES` output for the second-order EFA estimation. Using copy-paste, we construct the alternative second-order EFA estimation in Figure 6, which yields the same results as Figure 4 without random starting values and more advanced optimization parameter specifications. Notably, because we used the EFA as starting values, the factors are permuted as intended and all factors appear with a positive sign. When using random starting values, this will not necessarily occur. It should be noted that Mplus EFA estimation is fine-tuned so that larger loadings are always positive, and factors are ordered according to the order of the indicators they load on—that is, the factor ordering is not random. In addition, EFA estimation in Mplus benefits from specialized starting value procedures. None of these benefits are directly available for second-order EFA if we use random starting values. An additional benefit of using Figure 6 over Figure 4 is that the computation is faster, as a single likelihood optimization is performed. When using random starting values, separate likelihood optimizations are performed for each set of random starting values. Ultimately, Figure 6 is more likely to converge than Figure 4, as we do not have precise understanding of how many random starting values must be used to reach a converging model.

Figure 7 contains the results for the second-order EFA obtained with Figure 6. Figure 8 contains the results of the standard oblique EFA for comparison. The loading structure between the two models is largely unchanged. EFA and second-order EFA have the same number of significant cross-loadings and the same number of cross-loadings greater than 0.2 in absolute value. This implies that the second-order EFA model is suitable for the Holzinger–Swineford data and possibly provides a more practical interpretation of a general ability factor and specific domain factors. The EFA and the second-order EFA have the same number of free parameters (although not the same number of all parameters), log-likelihood values and chi-square test of fit as expected.

For comparison, we have also included the oblique bi-factor EFA results (the orthogonal bi-factor EFA results are very similar). Figure 9 contains the results for the bi-factor EFA with 4 factors and Figure 10 contains the results for 5 factors. The 4 factor model in Figure 9 is equivalent to the Figure 7 and Figure 8 results in terms of data fit. These are different rotations of the same unrotated EFA model with 4 factors. We see that with the bi-factor rotation, the spatial factor is lost. The first four indicators that were the primary indicators for the spatial factor now load only on the general factor. Using 5 factors, which in principle would allow space for the general factor to be separated from the specific factors, also does not recover the spatial factor. Here the last fifth factor does not have any significant loadings. Clearly, the bi-factor EFA model failed to support the substantive theory for this example.

Figure 4: Second order EFA input file for Holzinger and Swineford example

```
data: file is H-S Combined.txt;

variable:

names = id female grade agey agem school
visual cubes paper flags general paragrap
sentence wordc wordm addition code counting straight wordr
numberr figurer object numberf figurew deduct
numeric problemr series arithmet;

usev = visual-figurew;

define:   standardize visual-figurew;

analysis:
      estimator = mlr;
      iter = 10000; conv = 0.000001; starts = 50;

model:
      spatial verbal speed memory by visual-figurew*(a1-a76);
      spatial-memory (v1-v4);
      f0 by spatial-memory* (l1-l4); 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);
```

Figure 5: Using EFA estimation to obtain starting values for second order EFA

```
analysis:
      estimator = mlr;

model:
      spatial verbal speed memory by visual-figurew(*1);
      spatial verbal speed memory by visual-figurew(a1-a76);

output: svalues;
```

Figure 6: Second order EFA input file with starting values

```
analysis:
    estimator = mlr;

model:
   spatial BY visual*0.62097 (a1);
   spatial BY cubes*0.51635 (a2);
   spatial BY paper*0.46566 (a3);
   spatial BY flags*0.63592 (a4);
   ...
   memory BY figurer*0.45286 (a73);
   memory BY object*0.52555 (a74);
   memory BY numberf*0.39735 (a75);
   memory BY figurew*0.30527 (a76);
   spatial-memory (v1-v4);
   f0 by spatial-memory* (l1-l4); 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);
```

Figure 7: Second order EFA results for Holzinger and Swineford

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*

Figure 8: EFA results for Holzinger and Swineford

GEOMIN ROTATED LOADINGS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
VISUAL	0.621*	0.143*	0.032	0.059
CUBES	0.516*	0.037	-0.107	-0.013
PAPER	0.466*	0.090	0.014	-0.067
FLAGS	0.636*	-0.107	0.032	0.119
GENERAL	-0.016	0.846*	0.045	-0.079
PARAGRAP	0.010	0.803*	-0.006	0.076
SENTENCE	-0.056	0.909*	-0.005	-0.053
WORDC	0.080	0.696*	0.024	0.044
WORDM	0.069	0.818*	-0.032	0.035
ADDITION	-0.217*	0.025	0.764*	0.040
CODE	0.036	0.180*	0.541*	0.149*
COUNTING	0.115	-0.033	0.681*	-0.092
STRAIGHT	0.360*	0.004	0.505*	-0.042
WORDR	-0.037	0.087	-0.044	0.654*
NUMBERR	0.094	-0.120	-0.018	0.587*
FIGURER	0.325*	0.045	0.008	0.453*
OBJECT	-0.138	-0.026	0.325*	0.526*
NUMBERF	0.099	0.015	0.183*	0.397*
FIGUREW	0.087	0.178*	0.056	0.305*

GEOMIN FACTOR CORRELATIONS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
Spatial	1.000			
Verbal	0.356*	1.000		
Speed	0.303*	0.267*	1.000	
Memory	0.248*	0.215*	0.261*	1.000

Figure 9: Bi-factor EFA results for Holzinger and Swineford with 4 factors

ROTATED LOADINGS (* significant at 5% level)

	General	Verbal	Speed	Memory
VISUAL	0.709*	0.028	-0.063	-0.050
CUBES	0.459*	-0.044	-0.172*	-0.089
PAPER	0.470*	0.010	-0.058	-0.137
FLAGS	0.621*	-0.190*	-0.063	0.002
GENERAL	0.406*	0.730*	0.036	-0.074
PARAGRAP	0.452*	0.685*	-0.011	0.065
SENTENCE	0.389*	0.790*	-0.003	-0.041
WORDC	0.465*	0.584*	0.006	0.023
WORDM	0.487*	0.691*	-0.043	0.019
ADDITION	0.178*	0.046	0.713*	0.023
CODE	0.444*	0.141*	0.478*	0.094
COUNTING	0.372*	-0.049	0.588*	-0.145
STRAIGHT	0.564*	-0.053	0.396*	-0.128
WORDR	0.270*	0.061	-0.022	0.598*
NUMBERR	0.273*	-0.134*	-0.018	0.515*
FIGURER	0.533*	-0.023	-0.033	0.354*
OBJECT	0.232*	-0.020	0.320*	0.475*
NUMBERF	0.359*	-0.015	0.156*	0.330*
FIGUREW	0.332*	0.130*	0.041	0.257*

Figure 10: Bi-factor EFA results for Holzinger and Swineford with 5 factors

ROTATED LOADINGS (* significant at 5% level)

	General	Verbal	Speed	Memory	Spatial
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

5 Adding covariates

A benefit of a single second-order factor is that relations to other variables such as covariates are greatly simplified to concern only one factor. We illustrate this with the Holzinger and Swineford data where there are three covariates that can be used as predictors: grade, school, and gender. We consider four models. Model M1 is the ESEM model where all factors are regressed on all covariates. Model M2 is the second-order EFA with the second-order factor regressed on the covariates. In the most typical scenario, this is the model we want to consider. As it happens, however, real data examples do deviate from typical examples. If we only consider grade and gender as covariates, model M2 is easy to estimate and the analysis ends there. The school covariate, however, complicates the situation as it has an unusually strong direct effect on the verbal abilities factor. Thus the school covariate's indirect effect on the verbal factor via the general factor is therefore not sufficient. Estimating the M2 model with the school covariate doesn't converge because of that. We proceed with analyzing this more complex situation with all three covariates as it offers an opportunity to discuss a variety of different aspects of the second-order EFA as well as the power and complexities of the PSEM methodology.

Model M3 is the second-order EFA with all first- and second-order factors regressed on the covariates. ALF priors are added for the first-order factor regression coefficients. Model M3 is used to ensure that any possible direct effects from the covariates are included. The model uses a key building block of the PSEM framework (see Section 4.3 in Asparouhov and Muthén, 2024). Model M4 is the reduced M3 model, i.e., only statistically significant (as determined by M3) direct effects from the covariates to the first-order factors are included.

Model M3 is equivalent to model M1 in terms of data fit. Model M2 is not equivalent to model M1. Without the covariates, M2 and M1 are equivalent. However, when the second-order factor is used for modeling, in terms of being a predictor or being predicted, the two EFA models diverge. In our example, with 3 covariates and 4 first-order factors, model M1 has 9 more parameters than model M2. Model M4 is also not equivalent to M1, M2, or M3. M4 is nested above model M2 and is nested within M1 and M3.

For completeness, we also define model M3 in equation form

$$Y = \nu + \Lambda_1 F + \varepsilon, \tag{6}$$

$$F = B_1 X + \Lambda_2 \eta + \xi, \tag{7}$$

$$\eta = B_2 X + \zeta, \tag{8}$$

$$\varepsilon \sim N(0, \theta), \xi_i \sim N(0, 1 - \lambda_{2i}^2), \zeta \sim N(0, 1), \tag{9}$$

$$\Lambda_1 \sim Geomin(4, 0.1), B_1 \sim ALF(0, 1) \tag{10}$$

The last equation constitutes the definition of the penalty/prior for the M3 PSEM model.

Figure 11 contains the model statements for all four models. In our empirical example, the estimation of models M1, M3, and M4 converged, while model M2 did not, even when using EFA loadings as starting values. Model M3 found a strong direct effect from SCHOOL to VERBAL. This direct effect is the only direct effect in M3 that is statistically significant. There are two schools in this data: Grant–White and Pasteur. Students from the Grant–White school came from homes where the parents were mostly American born, whereas students from the Pasteur school came largely from working-class parents of whom many were foreign born and used their native language at home. This background information is the explanation for the strong direct effect from the school covariate to the verbal factor.

Model M4 is almost identical to model M2 but includes this one direct effect. Models M1 and M3 yield identical model fit. Model M4, however, yields the best BIC value, and when tested against M1 and M3, the more restricted model M4 is not rejected with an LRT p-value of 0.15. Thus, we conclude that model M4 provides the most complete explanation for the effect of the covariates on the factors. Figure 12 shows the covariate effects obtained with M1 and M4. The 12 parameters estimated with M1 are now summarized with 4 parameters in M4. In M1, the effect of the FEMALE predictor is insignificant for all factors and different in signs. In M4, this effect is summarized with one parameter decisively pointing towards no FEMALE effect on the general factor (and therefore on the domain factors). In M1, the effect of GRADE is uniformly positive and significant on all four factors. This is summarized in M4 as one significant effect on the general factor. The average Z-test score for the GRADE effects across the 4 factors in M1 is less than 4, while in M4, the Z-test score for the effect is more than 6. Here again we see the much more decisive conclusion that can be obtained with model M4, which accumulates information across the factors, and the improvement in the power of the model to detect significant effects. In M1, the effect of the SCHOOL covariate is completely different from the patterns of the first two covariates: significantly positive for one factor, significantly negative for another factor and not significant for the other two factors. In M4, this is summarized with one marginally significant effect on the general factor and a very strong direct effect in the opposite direction for the verbal factor. We conclude that the second-order EFA offers a superior summary of the covariate effects on the factors when compared to the general ESEM/EFA analysis with factor specific effect. It yields a more parsimonious model with more accurate results and more power to detect significance.

It is also interesting to note that the first order loadings for M4 are very close to those reported in Figure 7 for the model without the covariates. This is evidence that direct effects from the covariates to the indicators, beyond the one direct effects to the first-order factor, are not needed. Note however that the PSEM methodology can similarly be used to explore also direct effects to indicators, beyond the effects to the second order factors and beyond any potential direct effects to the first order factors. One such example is discussed in Section 4 in Asparouhov and Muthén (2025).

If we add covariates in a bi-factor ESEM model we get the general factor covariate effects separated from the specific factor covariate effects. This model directly corresponds to second-order EFA model M3, which also allows separation of covariates effect for the general abilities feature and the specific domain features. Similar construction exists also for the situation when the general and specific factors in a bi-factor model are used as predictors of other variables, see Gustafsson and Balke (1993). The first and second order factors in the second order EFA model can also be used as predictors for other variables. ALF priors must be added for the regression coefficients for the first order factors here as well for identification purposes. Additionally, if we want the predictors to be independent as in the orthogonal bi-factor EFA, the first order factors F must be replaced by their residuals ξ as in the RSEM (residual structural equation modeling) framework, see Asparouhov and Muthén (2021).

The degrees of freedom in PSEM models is a complex concept. The quantity is estimated numerically (see Asparouhov and Muthén, 2024). In many cases it is easy to understand, and in some cases it is not. The EFA model and the second-order EFA model have 108 free parameters. The EFA model has 120 parameters, and the geomin rotation eliminates (helps identify as constrained parameters) 12 as dependent parameters. The second-order EFA has 122 parameters; model constraints eliminate 4 as dependent parameters, and the geomin rotation eliminates an additional 10 as dependent parameters. Thus, the geomin rotation eliminates (helps identify) parameters differently for different models. Of course, this is already visible in standard EFA when comparing oblique and orthogonal rotation. In oblique EFA, the rotation identifies additionally all factor correlations.

Models M1 and M3 both have 120 parameters, as there are an additional 12 regression coefficients that are free parameters. Model M4, however, has 117 parameters, which is not as easy to understand. Here we have added just 4 regression parameters to the EFA model, but the number of parameters is not 112 as simple arithmetic would suggest. In reality, the restrictions on the regression coefficients (8 regression coefficients are fixed/constrained) can now act as a partial native model rotation. Native model rotation is a novel concept that has been discussed in Section 6.7 in Asparouhov and Muthén (2024). It is not a full native model rotation because without the geomin penalty, the model is not identified. In total, model M4 has 126 parameters (4 additional regression coefficients to the second-order EFA's 122 parameters). Four of these parameters are eliminated as dependent with model constraints. That means that the geomin rotation helps identify 5 additional parameters as dependent, not 10 as in the model without the covariates. The restriction on the regression coefficients provides part of the rotation, and part is provided by geomin. An exact explanation for how the two rotations combine to identify parameters is not easy to provide in complex models such as this one. Such an explanation is also not needed. We can simply rely on the numerical procedure responsible for counting the number of parameters, which is based on the rank of the information matrix. This is automatically provided in the Mplus output.

Figure 11: Second order EFA with covariates

```
**** Model M1 ****
model:
      spatial verbal speed memory by visual-figurew(*1);
      spatial verbal speed memory on female grade school;


**** Model M2 ****

  model:
      spatial verbal speed memory by visual-figurew*(a1-a76);
      spatial-memory (v1-v4);
      f0 by spatial-memory*(l1-l4); f0@1;
      f0 on female grade school;

  model constraint: DO(#,1,4) v#=1-l#*l#;

  model priors: a1-a76~Geomin(4,0.1);


**** Model M3 ****

  model:
      spatial verbal speed memory by visual-figurew*(a1-a76);
      spatial-memory (v1-v4);
      f0 by spatial-memory*(l1-l4); f0@1;
      f0 on female grade school;
      spatial-memory on female grade school (b1-b12);

  model constraint: DO(#,1,4) v#=1-l#*l#;

  model priors: a1-a76~Geomin(4,0.1); b1-b12~ALF(0,1);


**** Model M4 ****

  model:
      spatial verbal speed memory by visual-figurew*(a1-a76);
      spatial-memory (v1-v4);
      f0 by spatial-memory*(l1-l4); f0@1;
      f0 on female grade school;
      verbal on school;

  model constraint: DO(#,1,4) v#=1-l#*l#;

  model priors: a1-a76~Geomin(4,0.1);
```

Figure 12: Results for second order EFA with covariates

Model M1 (ESEM)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SPATIAL ON				
FEMALE	0.187	0.144	1.299	0.194
GRADE	0.478	0.154	3.103	0.002
SCHOOL	0.016	0.174	0.090	0.928
VERBAL ON				
FEMALE	0.121	0.123	0.980	0.327
GRADE	0.531	0.121	4.393	0.000
SCHOOL	-0.800	0.126	-6.346	0.000
SPEED ON				
FEMALE	-0.143	0.148	-0.965	0.334
GRADE	0.920	0.167	5.504	0.000
SCHOOL	0.511	0.189	2.700	0.007
MEMORY ON				
FEMALE	-0.105	0.151	-0.693	0.488
GRADE	0.400	0.159	2.514	0.012
SCHOOL	-0.095	0.230	-0.412	0.680

Model M4 (Second-order EFA with covariates and one direct efefct)

F0 ON				
FEMALE	-0.048	0.164	-0.296	0.767
GRADE	1.190	0.193	6.176	0.000
SCHOOL	0.508	0.219	2.322	0.020
VERBAL ON				
SCHOOL	-1.041	0.151	-6.894	0.000

6 Conclusion

As we have gained more experience with the HEFA/second-order EFA model, it is now necessary to revisit, update, and correct several aspects of this modeling technique: the estimation, the interpretation, and general expectations of what the modeling can do. It appears that with these additions, the model can truly take its rightful place as a serious alternative to bi-factor EFA modeling and general EFA models with substantial factor correlations that can clearly benefit from second-order factor modeling.

We have promoted here the view that the HEFA model's natural place is the PSEM framework, and not the ESEM framework via the EWC estimation of Morin et al. (2016) and Morin and Asparouhov (2018). Nevertheless, currently the HEFA analysis is primarily being conducted with the EWC approach. Clearly, further practical applications are needed to compare the two methods, as well as simulation studies. Our opinion is that the PSEM estimation is easier to set up and master than the EWC approach; however, others might disagree. There is no disagreement, however, that we now have much better tools to pursue the HEFA model, even if the two different approaches are used simultaneously.

The focus of this paper was to provide an improved estimation for the case where the first-order model is an EFA model and there are one or more second order EFA or CFA factors. However, the concepts we discussed here may apply to some other scenarios. Asparouhov and Muthén (2025) discuss the model where the first-order analysis is CFA, while the second order analysis is EFA. For that model estimation, conceptually a similar issue arises: the geomin rotation function is affected by how the first order factor scales are set. The solution used in Asparouhov and Muthén (2025) is to set the scale of the first order factors by fixing one CFA loading parameter to 1. That way the scale of the first order factor is tied to the scale of the indicator and it would be difficult for the geomin optimization to manipulate the second order loadings into zeros without a substantial loss of fit. Nevertheless, some additional caution is advised. If the scale is set using a poor indicator, for example, the first order factor variance will be small and the second order loadings will be near zero, which will eliminate the first order factor as a source of information for the second order EFA analysis.

References

[1] Asparouhov T, Muthen B. (2009). Exploratory structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 16, 397– 438.

[2] Asparouhov, T., & Muthén, B. (2021). Residual structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, 30, 1-31.

[3] Asparouhov, T., & Muthén, B. (2024). Penalized Structural Equation Models. Structural Equation Modeling: A Multidisciplinary Journal, 31, 429–454.

[4] Asparouhov, T., & Muthén, B. (2025) Methodological Advances with Penalized Structural Equation Models. Structural Equation Modeling: A Multidisciplinary Journal, 32, 688-716.

[5] Gustafsson, J.E., Balke, G. (1993). General and specific abilities as predictors of school achievement. Multivariate Behavioral Research, 28, 407-434

[6] Jennrich R.I., Bentler P.M. (2011). Exploratory bifactor analysis. Psychometrika, 76, 537-49.

[7] Jennrich R.I., Bentler P.M. (2012). Exploratory bifactor analysis: The oblique case. Psychometrika, 77, 442-54.

[8] Harman, H.H. (1967). Modern factor analysis. Second edition. Chicago, IL: The University of Chicago Press.

[9] Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. Psychometrika, 2, 41–54.

[10] 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.

[11] Morin, A.J.S., & Asparouhov, T. (2018). Estimation of a hierarchical Exploratory Structural Equation Model (ESEM) using ESEM-within-CFA. Montreal, QC: Substantive Methodological Synergy Research Laboratory. https://www.statmodel.com/download/Webnote%20-%20Hierarchical%20Exploratory%20Structural%20Equation%20Model.pdf

[12] Morin, A. J. S., Arens, A., & Marsh, H. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. Structural Equation Modeling, 23, 116-139.

[13] Muthén, B. & Asparouhov, T. (2012) Bayesian structural equation modeling: a more flexible representation of substantive theory. Psychological Methods 17, 313-335.

[14] Reise S.P. (2012) The rediscovery of bifactor measurement models. Multivariate Behavioral Research, 47, 667–696.