A Unification of Second-Order and Bi-Factor EFA

Tihomir Asparouhov and Bengt Muthén

Version 2

Mplus

April 22, 2026

Abstract

As empirical applications and methodological research on second-order exploratory factor analysis have matured, important refinements to estimation procedures, interpretive frameworks, and theoretical expectations have emerged. These advances establish second-order EFA as a rigorous alternative to bi-factor EFA and to conventional EFA models characterized by substantial factor correlations. We further demonstrate that bi-factor EFA models constitute a reparameterization of second-order EFA with direct effects, clarifying the formal relationship between these approaches. The unification of bi-factor and second-order EFA models reveals a novel rotation approach for bi-factor solutions that outperforms bi-geomin in simulation studies and yields more interpretable factor solutions in empirical studies.

1 Introduction

Exploratory factor analysis (EFA) is used to explain correlations among observed variables through underlying latent constructs. In oblique EFA, these latent constructs are themselves correlated, and such correlations are often substantial. These factor correlations, in turn, can be accounted for by an additional latent variable — commonly interpreted as a general latent construct that exerts a broad influence on all observed outcomes. Two EFA frameworks exist for modeling such a general construct: the bifactor model and the second-order factor model. While these two models have been treated as distinct, this paper shows that they can be viewed in a unified framework. The paper describes recent methodological advances for both models and introduces a new model: the direct-effect second-order EFA (DSEFA). DSEFA provides a synthesis of bifactor and second-order approaches, and can be interpreted within either framework, thereby offering insight into the relationship between the two models.

Asparouhov and Muthén (2024a, 2025) developed Second-order Exploratory Factor Analysis (SEFA) within the Penalized Structural Equation Modeling (PSEM) framework. The model specifies a two-level hierarchy: observed indicators load on first-order factors, which in turn load on second-order factors. Both levels accommodate exploratory or confirmatory specifications:

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

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

where Λ_1 and Λ_2 are loading matrices (unconstrained in EFA, structured in CFA), F contains first-order factors, and η contains second-order factors. The model is also known as hierarchical EFA (HEFA).

The DSEFA model expands on the SEFA model by introducing direct effects from the second order factors to the indicators, i.e., equation (1) is replaced by

$$Y = \nu + \Lambda_1 F + \Lambda_3 \eta + \varepsilon. \tag{3}$$

The directed effects parameters Λ_3 are unconstrained exploratory parameters available in the PSEM framework as identifiable parameters, see Section 4.3 in Asparouhov and Muthén (2024a). In standard SEM modeling, adding unconstrained Λ_3 to the model yields an unidentified model, but in the PSEM framework the parameters are identifiable through the process of aligning the effects of η on Y through the factors F and the direct effects. If Λ_2 is set to 0, the DSEFA model becomes identical to the bi-factor model and thus it can be viewed as a fusion of SEFA and bifactor EFA.

This paper provides corrections and clarifications to Section 6.2 of Asparouhov and Muthén (2024a), which describes SEFA with exploratory first-order analysis. Accumulated practical experience with this model has led to three key refinements: improved PSEM-based estimation procedures, enhanced interpretive guidelines, and clearer positioning relative to alternative models such as bi-factor EFA via the DSEFA model. We address each aspect in detail below. The most commonly estimated SEFA features multiple exploratory first-order factors and a single second-order factor. While we focus primarily on this case, our conclusions extend to models with multiple second-order factors, exploratory second-order structures, and specifications incorporating observed second-order indicators. This work also demonstrates the flexibility of the PSEM framework in Mplus.

Morin et al. (2016) introduced the SEFA model using a two-stage estimation procedure (the EWC approach) within the ESEM framework of Asparouhov and Muthén (2009). Notably, this method also required subsequent revision for reasons similar to those necessitating the present PSEM refinements (see Morin and Asparouhov, 2018). The PSEM-based approach offers several advantages over the EWC method: single-stage estimation, simpler implementation, and reduced susceptibility to estimation errors. The methodological refinements presented here should therefore facilitate more reliable applications of the SEFA model.

In Mplus version 9.1, SEFA and DSEFA models are estimated directly by specifying SEFA and DSEFA as the EFA rotation. We illustrate these methods with empirical examples and simulation studies. Mplus input files are also provided. A key finding of our simulation studies is that the DSEFA rotation outperforms the bi-geomin rotation of Jennrich and Bentler (2011) for bi-factor models.

The paper is organized as follows. Section 2 introduces improvements to SEFA estimation and illustrates the model with a simulation study. Section 3 addresses SEFA interpretation and model fit, and situates second-order rotations within the classical orthogonal/oblique EFA framework. Section 4 establishes the relationship between second-order and bi-factor EFA and introduces the direct-effect second-order EFA (DSEFA) model. Section 5 presents extensive simulation studies comparing eight EFA methods under two distinct data-generating scenarios, demonstrating the superior factor recovery of SEFA and DSEFA. Section 6 illustrates the new methods using the Holzinger and Swineford (1937) dataset and examines the challenging problem of rotation selection, where traditional fit-based model selection criteria are inadequate. Section 7 extends the proposed EFA framework to general exploratory structural models by incorporating covariates, broadening the scope of ESEM. Section 8 concludes.

2 Estimation of the SEFA model

The SEFA model is estimated via PSEM using Geomin(Λ_1) as the penalty function, where Geomin denotes the rotation function for the geomin rotation criterion

$$\text{Geomin}(\Lambda) = \sum_{i=1}^{p} \left(\prod_{j=1}^{m} (\lambda_{ij}^2 + \epsilon) \right)^{\frac{1}{m}}, \tag{4}$$

where p is the number of indicators, m is the number of factors and ϵ is a small smoothing constant typically in the range of 0.0001 to 0.01. Figure 13 in the supplemental materials of Asparouhov and Muthén (2024a) demonstrates the implementation. Simulation studies showed satisfactory performance; however, subsequent practical applications revealed convergence difficulties, particularly with smaller samples. For instance, the Figure 13 specification achieves 100% convergence at $N = 2000$, but only 75% at $N = 500$ and 46% at $N = 300$. We propose a modified estimation approach that achieves 100% convergence across all three sample sizes. The key modification concerns first-order factor scaling. The original specification fixes residual variance at unity: $\text{Var}(\xi) = 1$. We instead constrain total factor variance: $\text{Var}(F) = 1$, paralleling the scaling convention in single-order EFA. This simple change resolves the convergence issues.

When scale is identified solely through residual variance, the penalized likelihood exhibits multiple local maxima competing with the true solution. Because Geomin penalization favors near-zero loadings, pathological local maxima emerge through the following mechanism: for any first-order factor, dividing its first-order loadings by a large constant c while multiplying its second-order loading by c preserves the overall covariance structure but causes the second-order factor to collapse onto that first-order factor. Although this transformation degrades data fit, the penalty function improvement compensates for the likelihood loss. These spurious maxima may occasionally represent global optima, but more commonly, they attract the optimization algorithm away from the true solution. With m first-order factors, the optimizer confronts m such competing local maxima. Constraining total factor variance ($\text{Var}(F) = 1$) eliminates this indeterminacy by decoupling the two loading matrices, preventing any multiplicative constant from propagating between hierarchical levels.

The first-order factor variances are functions of parameters rather than model parameters

$$\text{Var}(F_i) = \lambda_{2i}^2 + \text{Var}(\xi_i). \tag{5}$$

Thus, the SEFA model estimation is a constrained maximum likelihood optimization where the first order factor residual variances are directly determined by the second order factors

$$\text{Var}(\xi_i) = 1 - \lambda_{2i}^2. \tag{6}$$

2.1 SEFA simulation study

Figure 1 illustrates the SEFA implementation in Mplus with a simulation study based on four first-order and one second-order factors. The MODEL POPULATION: section of the input file contains the exact parameter values used for the data generation. The values for the residual variances are chosen to satisfy the constraint (6). The MODEL: section specifies the estimated model. Starting values are given so that proper confidence interval coverage rates are computed. Mplus will use these starting values as the true parameter values. Alternatively, MODEL COVERAGE: can be specified for that purpose and the starting values can be removed from the MODEL: section. This can be useful to study the ability of the algorithm to reach convergence without good starting values. The ROTATION=SEFA option automatically specifies the GEOMIN prior for Λ_1 and the parameter constraint (6). Figure 2 presents the results for selected first-order loadings and all second-order loadings. Results demonstrate adequate parameter recovery at $N = 1000$. As shown in the next section, the chi-square test for SEFA with m first-order factors is identical to that for standard m-factor EFA. Here, the chi-square test of fit average is 117 and with 116 degrees of freedom yields a 7% rejection rate.

We also conduct this simulation study with sample sizes $N = 100, 200, 300,$ and 500 to evaluate the performance of the estimator for smaller sample sizes. Overall, the results are very similar. As sample size decreases, the bias of the second-order loadings increases slightly. For $N = 1000$, the bias of the first second-order loading is 0.01, while at $N = 300$ it is 0.03. This is expected for most maximum-likelihood-based procedures, where bias is guaranteed to be near zero only for sufficiently large samples. Convergence rates remain at 100%, although at $N = 100$ the default number of iterations is not enough and the `ANALYSIS` option `ITER=10000` is needed to increase convergence from 91% to 100%. Using the `CONVERGENCE=STRICT` option, we can also discard Heywood cases. Inadmissible solutions are found only for $N = 100$ and $N = 200$.

It is instructive to compare inadmissible solutions between SEFA and EFA. Heywood cases found in the second-order model arise because the second-order factor model is a poor fit for the first-order factor correlations. We can therefore expect that these will not occur in standard EFA, where the factor correlation matrix, being a rotation of the identity matrix, is always admissible. Consequently, Heywood cases arising in the second-order model are of primary concern, as they are attributable exclusively to SEFA. Heywood cases found in the first-order model should presumably be identical across all EFA methods, as they reduce to negative residual variances for the observed variables, which are not affected by the rotation method. In practice, however, they are not exactly identical here, because EFA estimates the unrotated model to determine these residual variances while SEFA estimates the rotated model directly. With a large number of random starting values, these will converge to the same result.

For $N = 200$, there are 3% second-order Heywood cases in SEFA and 8% first-order Heywood cases, which is 1% more than in EFA. For $N = 100$, there are 7% second-order Heywood cases in SEFA and 21% first-order Heywood cases, which is 6% more than in EFA. We conclude that in small samples, SEFA estimation carries a slightly elevated risk of inadmissible solutions due to increased Heywood cases. This concern is further tempered, however, by the fact that the average four-factor EFA BIC falls below that of the three-factor solution only at $N \geq 300$. In small-sample settings, therefore, the primary obstacle is likely establishing statistical support for the required number of EFA factors — SEFA requires at least three — rather than Heywood cases or finite-sample bias.

Figure 3 shows the much simpler input syntax for empirical applications where starting values are typically unavailable and are not needed. The SEFA rotation option can be used to alter the weight of the GEOMIN penalty. This might be necessary if the variables are not standardized or the sample size is very large or small.

Figure 4 illustrates the manual setup for SEFA model estimation, in which the constraint (6) is specified in `MODEL CONSTRAINT` and the geomin rotation is imposed as a penalty function in `MODEL PRIOR`. With this approach, it is essential to include the option `STARTS=50`, as the automatic starting-value procedure used in standard EFA is not invoked. This manual setup can be applied to models for which `ROTATION=SEFA` is not yet available, such as mixture or multilevel models.

Asparouhov and Muthén (2024a) reported that SEFA estimation was more challenging than standard or bi-factor EFA. This difficulty, however, was an artifact of the residual-variance scaling approach ($\text{Var}(\xi) = 1$) employed at that time. With the revised total-variance constraint ($\text{Var}(F) = 1$), SEFA estimation is comparable in difficulty to standard EFA models. This improved convergence holds for both simulated and empirical data.

Figure 1: SEFA simulation study

```
MONTECARLO:
NAMES = y1-y20;
NOBSERVATIONS = 1000;
NREPS = 100;

analysis:  rotation=sefa;

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(*1);
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 (*1);
f3 BY y1-y10*0 y11*0.7 y12*1 y13-y14*0.4 y15*0.4 y16-y20*0 (*1);
f4 BY y1-y14*0 y15*0.6 y16*0.7 y17*1 y18*0.5 y19*1 y20*0.5 (*1);
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;
```

Figure 2: SEFA simulation study results

		ESTIMATES		S. E.	M. S. E.	95%	% Sig
	Population	Average	Std. Dev.	Average		Cover	Coeff
F1 BY							
Y1	0.700	0.6953	0.0441	0.0442	0.0019	0.930	1.000
Y2	1.300	1.2936	0.0550	0.0578	0.0030	0.950	1.000
Y3	0.800	0.7946	0.0421	0.0460	0.0018	0.970	1.000
Y4	0.800	0.7983	0.0483	0.0458	0.0023	0.920	1.000
Y5	0.300	0.3053	0.0402	0.0461	0.0016	0.950	1.000
Y6	0.000	0.0089	0.0343	0.0401	0.0012	0.980	0.020
Y7	0.000	-0.0015	0.0395	0.0405	0.0015	0.950	0.050
Y8	0.000	0.0024	0.0345	0.0411	0.0012	0.990	0.010
Y9	0.000	0.0023	0.0370	0.0427	0.0014	0.990	0.010
Y10	0.000	0.0049	0.0407	0.0402	0.0017	0.960	0.040
Y11	0.000	0.0007	0.0403	0.0425	0.0016	0.960	0.040
Y12	0.000	0.0018	0.0374	0.0435	0.0014	0.990	0.010
Y13	0.000	0.0039	0.0363	0.0402	0.0013	0.980	0.020
Y14	0.000	0.0033	0.0414	0.0401	0.0017	0.950	0.050
Y15	0.000	0.0009	0.0240	0.0299	0.0006	1.000	0.000
Y16	0.000	0.0019	0.0381	0.0413	0.0014	0.970	0.030
Y17	0.000	0.0078	0.0356	0.0420	0.0013	1.000	0.000
Y18	0.000	-0.0004	0.0369	0.0399	0.0013	0.970	0.030
Y19	0.000	0.0025	0.0395	0.0436	0.0016	1.000	0.000
Y20	0.000	0.0038	0.0439	0.0403	0.0019	0.930	0.070
F0 BY							
F1	0.400	0.3902	0.0601	0.0779	0.0037	0.990	0.990
F2	0.500	0.4867	0.0615	0.0824	0.0039	1.000	1.000
F3	0.500	0.4887	0.0742	0.0915	0.0056	0.970	0.980
F4	0.600	0.5923	0.0692	0.0893	0.0048	0.980	1.000

Figure 3: SEFA input file for real data analysis

```
variable: names = y1-y20;

data: file=1.dat;

model:
f1-f4 BY y1-y20(*1);
f0 by f1-f4;

analysis: rotation=sefa;
```

Figure 4: Alternative SEFA input file for real data analysis manual setup

```
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 of the SEFA model

SEFA with m_1 first-order factors and m_2 second-order factors has the same fit as EFA with m_1 factors, regardless of whether the second-order model is an EFA or a CFA model. This can be shown as follows. SEFA 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, SEFA with m_1 first-order factors, in terms of log-likelihood, cannot exceed the log-likelihood of EFA with m_1 factors. Now we will show that the opposite is true: SEFA maximum likelihood 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 SEFA is unconstrained, this solution is among those that are considered in the SEFA 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, SEFA 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. This 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, 2024a, for details). The null model for SEFA 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 SEFA 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 to 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 concerns only the allowed rotations. The rotation in SEFA is neither oblique nor orthogonal. The rotations are a subset of the oblique rotations that conform to the second-order model. This also means that we cannot evaluate the second-order model using the data fit, just as we cannot choose between oblique or orthogonal EFA based on the data fit: they have identical data fit.

The advantage of oblique EFA over orthogonal EFA is that a simpler loading structure can be obtained. The disadvantage is that the factor correlation matrix is unrestricted and can be viewed as more complex. SEFA offers a compromise: a simpler loading structure than orthogonal EFA with a more parsimonious and interpretable first-order factor correlation structure based on the second-order factor model. The second-order factor can be interpreted as an inherent characteristic affecting all first-order factors. Furthermore, the second-order factor yields a parsimonious model when relating to covariates or when serving as a predictor of other variables.

SEFA, like orthogonal EFA, compensates for its more restrictive factor correlation structure by allowing a few additional cross-loadings (if needed) compared to oblique EFA. SEFA lies between orthogonal and oblique EFA on the scale of loading structure simplicity for the indicators. It also falls between the two traditional EFA models in terms of factor correlation simplicity.

The main question that needs to be answered is when SEFA should be preferred over traditional orthogonal and oblique EFA. If the SEFA model's first-order loading structure Λ_1 is better and more interpretable than the orthogonal EFA loading structure and is comparable to (not substantially worse than) the oblique model, it should be preferred and considered the best among the three models.

Consider the example in Figure 3. The data are generated using the simulation study in Figure 1. There are 22 first-order loadings that are nonzero. The orthogonal model yields a factor structure with 28 significant loadings. In contrast, SEFA and oblique EFA yield nearly identical loading structures: a total of 24 significant loadings (with both models finding 2 spurious cross-loadings). We conclude that restricting oblique EFA to a second-order factor model does not result in a worse loading structure. This provides evidence for the existence of a second-order factor explaining the factor correlations. In this example, the SEFA model should be preferred.

One question that naturally arises is whether the precision of the loading estimates (i.e., MSE) is affected by or benefits from a more restrictive factor correlation structure. The answer appears to be no, or at least any effect is not of substantive importance. The simulation study presented in Figure 1 yields identical loadings for SEFA and oblique EFA within each replication (and therefore identical MSE). The more parsimonious SEFA does not appear to meaningfully reduce MSE.

Oblique EFA may sometimes produce factor correlations that are too high to reliably distinguish between factors. SEFA 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 minima of the geomin optimization can be found in close proximity. SEFA has a more restrictive domain of rotations over which to optimize, which will likely 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 covariance matrix may lead to an alternative geomin minimum. Thus, we expect that another advantage of SEFA is the stability of the rotation.

4 Comparison of SEFA and bi-factor EFA

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, a two-stage estimation was proposed by Schmid and Leiman (1957). A single-stage estimation was established in Jennrich and Bentler (2011, 2012). More recently, a number of other estimation methods have been proposed. A comprehensive overview of the various estimation methods is available in Reise et al. (2023).

When equations (1) and (2) are combined, the SEFA model

$$Y = \nu + \Lambda_1 \xi + \Lambda_1 \Lambda_2 \eta + \varepsilon \tag{7}$$

resembles a bi-factor model. SEFA with m_1 first-order factors and m_2 second-order factors is equivalent to a factor model with $m_1 + m_2$ factors: ξ and η, which are independent. The ξ factors are also mutually independent. Note that the factors ξ are the 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, because the loading matrix for the general factors $\Lambda_1\Lambda_2$ is not unrestricted as in EFA. Most importantly, however, in terms of data fit, this SEFA model 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. It is simply a conclusion that follows from the fact that SEFA 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 the case of one general factor, SEFA with m_1 first-order factors and 1 second-order factor is equivalent in terms of log-likelihood and number of free parameters to the bi-factor EFA with $m_1 - 1$ specific and 1 general factor. On the other hand, equation (7) yields a bi-factor model with m_1 specific factors and 1 general factor. This discrepancy in the factor count may appear to be an obstacle to properly understanding the relationship between the two EFA models.

This discrepancy is also referred to as the rank deficiency in Waller (2018) and Reise et al. (2023), in connection with the estimation of bi-factor EFA models. Schmid and Leiman (1957) use the connection between SEFA and bi-factor EFA to construct an estimation method for bi-factor EFA with the caveat that it has the wrong number of factors. Subsequently, several methods have been proposed to deal with the rank deficiency (Reise et al., 2023). In fact, currently the estimation of bi-factor EFA is still very much linked to equation (7) and to the Schmid and Leiman (1957) method. A comprehensive simulation study conducted in Giordano and Waller (2020) shows that bi-factor EFA models are best estimated with methods based on Schmid and Leiman (1957) rather than the GPA method of Jennrich and Bentler

(2011-2012), which is in fact the most widely used method in practice. Failures of the GPA method have also been documented in Mansolf and Reise (2016), specifically near the parameter space where bi-factor EFA is collapsible to SEFA with one fewer factor.

The PSEM framework allows us to contribute two methods to bi-factor EFA estimation: the direct effect SEFA (DSEFA) and the Target-ALF method. We define these for the case of orthogonal bi-factor EFA where the specific factors are uncorrelated, but the methods naturally generalize to oblique bi-factor EFA as well. The DSEFA model also allows us to provide a direct relationship (reparameterization) between SEFA, DSEFA, and bi-factor EFA, as well as explain the rank deficiency in natural terms. Ultimately, DSEFA provides simple and accurate estimation for bi-factor EFA, without specifying targets and is easily generalizable to bi-factor models with more than one general factor.

4.1 DSEFA model

To formulate the DSEFA model, we first give the full definition of the SEFA model in the PSEM framework, assuming a single second-order factor and $m_1 - 1$ first-order exploratory factors F

$$Y = \nu + \Lambda_1 F + \varepsilon, \tag{8}$$

$$F = \Lambda_2 \eta + \xi, \tag{9}$$

$$\varepsilon \sim N(0, \theta), \xi \sim N(0, I - \Lambda_2 \Lambda_2^T) \tag{10}$$

$$\Lambda_1 \sim \text{Geomin}(m_1 - 1, w_1) \tag{11}$$

The SEFA model is equivalent to an EFA with $m_1 - 1$ factors in terms of data fit. The DSEFA model includes direct effects from the second-order factor to the indicators by adding the term $\Lambda_3 \eta$ in the first-order model

$$Y = \nu + \Lambda_1 F + \Lambda_3 \eta + \varepsilon, \tag{12}$$

$$F = \Lambda_2 \eta + \xi, \tag{13}$$

$$\varepsilon \sim N(0, \theta), \xi \sim N(0, I - \Lambda_2 \Lambda_2^T), \tag{14}$$

$$\Lambda_1 \sim \text{Geomin}(m_1 - 1, w_1), \Lambda_3 \sim \text{ALF}(0, w_2). \tag{15}$$

The DSEFA model is equivalent to an EFA with m_1 factors in terms of data fit. All direct effects in Λ_3 are given an ALF prior for identification purposes. Combining equations (12) and (13), we obtain the bi-factor EFA model

$$Y = \nu + \Lambda_1\xi + (\Lambda_1\Lambda_2 + \Lambda_3)\eta + \varepsilon, \tag{16}$$

where the general factor loadings in terms of the bi-factor EFA scale are

$$\Lambda_1\Lambda_2 + \Lambda_3 \tag{17}$$

while the specific loadings must be rescaled as

$$\Lambda_1\sqrt{I - \Lambda_2\Lambda_2^T} \tag{18}$$

to ensure specific factors are standardized. Unlike the SEFA model, DSEFA does not have a rank deficiency. The factor count, as well as data fit, log-likelihood, and free parameters, matches the bi-factor model with $m_1 - 1$ specific factors and 1 general factor. We conclude that DSEFA is a direct reparameterization of bi-factor EFA and can be used for its estimation, with one-step estimation and without specifying targets. Furthermore, the DSEFA model implements the Schmid and Leiman (1957) bi-factor EFA estimation method with a single-step direct estimation. The model also provides a clear connection between the factors in the second-order model and the factors in bi-factor EFA. The second-order factor in DSEFA is the same as the general factor in bi-factor EFA. The specific factors in bi-factor EFA are the standardized residuals of the first-order factors in DSEFA.

The DSEFA model has a very natural relationship to the SEFA model as well. The two are equivalent if the second-order factor does not have any direct effects on the indicators. If no direct effects exist, then the data can be fitted with one fewer factor. The DSEFA modeling approach, which provides a unifying framework connecting second-order and bi-factor modeling, can also be used with CFA models.

The DSEFA model also connects to the Jennrich and Bentler (2011-2012) estimation method, implemented in Mplus with the bi-geomin rotation. The difference between the two methods is that they use slightly different rotation criteria. In Jennrich and Bentler (2011-2012), the specific factor loadings are rotated to simplicity with the Geomin rotation function while the loadings of the general factor are left completely unconstrained. In DSEFA, the specific factors are treated the same way, but the general factor loadings also participate in the rotation and are constrained to be as close as possible to a linear combination of the specific factor loadings so as to minimize the need for direct effects. This naturally resolves the problems with the bi-geomin rotation documented in Mansolf and Reise (2016). For the simulation study in Figure 1, the bi-geomin rotation with 3 and 4 specific factors does not recover the general structure of the model. One of the specific factors is combined with the general factor in both cases, and certain indicators are without a specific factor. On the other hand, the DSEFA model yields results consistent with the data generation: there are no direct effects from the general factor to the indicators, and there are 4 specific factors.

Figure 5 shows how DSEFA model estimation is specified in Mplus. With this specification, Mplus produces standardized results for both the second order EFA model (12-13) as well as the equivalent bi-factor EFA model (16). The ROTATION=DSEFA option automatically sets the parameter constraints (14) and the priors (15). The manual setup for the same model is illustrated in Figure 6.

In practice, it is not unusual for bi-factor models to be presented where certain indicators are considered "general ability" indicators—that is, 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" construct in favor of mathematics. From that perspective, the SEFA model offers an alternative interpretation that some might find less biased and less controversial. In addition, the SEFA model structure accommodates "general ability" indicators more naturally. If such indicators are truly designed as "general ability" indicators, they belong in equation (9) and not in equation (8). Such indicators should be used directly as indicators for the general factor rather than in the rotation measuring domain-specific factors.

One disadvantage of the DSEFA method is that it uses two separate penalty weights: one for the rotation of the specific loadings and a second weight for the direct effects. Each of these must be deter-

Figure 5: DSEFA input file for real data analysis with 4 first order factors (Bi-factor EFA with 1 general and 4 specific factors)

```
variable: names=y1-y20;
data: file=1.dat;
model: f0 f1-f4 BY y1-y20(*1);
analysis: rotation=dsefa;
```

Figure 6: DSEFA input file for real data analysis manual setup

```
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;
    f0 by y1-y20*(d1-d20);

model constraint:
v1=1-L1*L1;
v2=1-L2*L2;
v3=1-L3*L3;
v4=1-L4*L4;

model priors: a1-a80~Geomin(4,0.1,0.01); d1-d20~ALF(0,1);

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

mined separately. The general strategy is to reduce each weight (i.e., increase the prior variance) until the log-likelihood stops improving. The final weight selection should reproduce the log-likelihood of the EFA with m_1 factors, which serves as the target value.

This selection process resembles sensitivity analysis in that models with different prior variances are compared, but it is considerably simpler. Rather than characterizing how results change across a range of values, the only quantity of interest is the point at which the log-likelihood plateaus — and that plateau value is known in advance, since it equals the log-likelihood of the null model (EFA with m_1 factors). The precise weight at which this plateau occurs is itself unimportant. Unlike regularized or Bayesian approaches, the penalty weight here is not intended to balance the prior against the information in the data. Its role is solely to ensure that, given the chosen convergence criteria, the log-likelihood is fully optimized, and that among all solutions achieving the optimal log-likelihood, the penalty is minimized as well.

In practice, this means that for nearly all PSEM applications, the sensitivity analysis reduces to estimating the model at prior variances of 1 and 0.1. Values outside this range are rarely needed, and when they are, successive powers of 10 in either direction generally suffice. For DSEFA specifically, bivariate sensitivity analysis over both weights simultaneously is not anticipated to be necessary; two sequential univariate analyses should be sufficient.

One natural sequencing is as follows. First, fix the direct effects prior to a very small value and vary the specific factor loading rotation weight until the log-likelihood plateaus at the value corresponding to EFA with $m_1 - 1$ factors. Then, holding the rotation weight fixed, increase the direct effects prior variance until the log-likelihood plateaus at the value corresponding to EFA with m_1 factors.

The DSEFA method can also be used to estimate oblique bi-factor EFA, as in the oblique bi-geomin method of Jennrich and Bentler (2012), but with two additional advantages. First, correlations can be restricted to the specific factors, thereby avoiding the unnecessary introduction of correlations between the general and specific factors, which are almost universally estimated to be zero in the oblique bi-geomin method. Second, all first-order (specific) factor correlations can be given ALF(0,1) priors as in Sections 4.5 and 4.6 of Asparouhov and Muthén (2024a). This further enriches the rotation criterion and reduces the burden placed on the geomin optimization, which in the bi-geomin method is responsible for not only simplifying the loading patterns but also identifying the specific factor correlations — subsequently leading to a somewhat poorly identified model with inflated standard errors. In DSEFA, adding ALF(0,1) priors for the specific factors amounts to accommodating any deficiency in the second-order factor model that may arise when simplifying the first-order loading matrix.

4.2 Bi-factor EFA with Target-ALF rotation

The second estimation method available in the PSEM framework for the estimation of bi-factor EFA is the Target-ALF method. This method is similar to the Target rotation method used for bi-factor EFA model estimation but instead of the sum of squares loss function (normal priors), it uses the ALF (square root) loss function. In most situations, the ALF loss function performs much better than the sum of squares and slightly better than the LASSO loss function (see Asparouhov and Muthén, 2024a). It is also shown in Asparouhov and Muthén (2024a) that Target-ALF rotation outperforms the traditional Target rotation in situations where small but nonzero cross-loadings are specified as targets—that is, the targets are incorrect. The Target-ALF rotation is generally safer to use than Target as it is not affected by incorrect targets. For bi-factor EFA, targets are given for the specific factors but the general factor loadings are left unconstrained. The limitations of the traditional Target rotation have also been noted in Moore et al. (2015), where an iterative procedure is suggested: after an initial run, incorrect targets are replaced with new targets until the targets no longer change. The Target-ALF procedure does not require such an iterative approach since it is not affected by incorrect target specification. A formal definition of the Target-ALF bi-factor EFA model is as follows. If F represents a vector of specific factors and η is the general factor, the model is given by

$$Y = \nu + \Lambda_1 F + \Lambda_2 \eta + \varepsilon, \tag{19}$$

$$\text{subset of } \Lambda_1 \sim \text{ALF}(0, w). \tag{20}$$

5 Simulation study comparing frameworks, models, rotation and number of factors

In this section, we illustrate several key concepts regarding bi-factor EFA, SEFA and DSEFA models. In particular, we demonstrate that the "distance" between SEFA and bi-factor EFA is a critical concept: the closer the two models are, the greater the advantage SEFA and DSEFA offer. We do not provide a precise definition of "distance" between the two models, but we offer the following heuristic explanation. A SEFA model with m factors naturally translates to a bi-factor EFA with $m + 1$ factors, as shown in equation (7). We consider a bi-factor EFA model with $m + 1$ factors to be close to a SEFA model with m factors if the general factor loadings approximate the form $\Lambda_1\Lambda_2$. Under the assumption of one general factor, this means that the general factor loadings are a linear combination of the specific factor loadings. If we further assume that Λ_1 does not include any cross-loadings, the two models would be considered close if the specific loadings of a particular factor are approximately proportional to the general factor loadings, with the coefficient of proportionality being the second-order loading. An alternative way to define closeness between SEFA and bi-factor EFA is that the corresponding DSEFA model has minimal direct effects. The concept of "distance" between the two models has been discussed in the literature under the names "proportionality" or "hierarchical" (Mansolf and Reise, 2016, 2017; Waller, 2018; Reise et al., 2023).

We demonstrate here that bi-factor EFA model estimation performs surprisingly poorly in a particular region of the total parameter space. The proportionality constraints are much easier to satisfy when there are fewer indicators per specific factor, which directly implies that bi-factor EFA is best estimated with a larger number of indicators per specific factor. A safe recommendation would be to have at least 10 indicators. Fewer than 5 indicators per specific factor is likely to be problematic. Between 5 and 10 indicators, potential proportionality problems should be investigated and simulation studies should be performed. There are some exceptions. If some indicators load only on the general factor, we do not anticipate difficulties. Similarly, if the general factor has a mix of high positive and negative loadings (corresponding to the same specific factor) while all specific loadings are positive, proportionalities are not present.

We conduct two simulation studies. In the first, we generate data using a 3-factor SEFA model. In the second, we generate data using a 4-factor bi-factor model that is sufficiently different from a 3-factor SEFA model. In both cases, we use 12 indicators with 4 indicators per specific factor. One small cross-loading is included.

Table 1: EFA methods (Sorted by MSE Performance), Data generated from a 3-factor SEFA

Method	Rotation	Model	Framework	Factors	DF	χ^2 Mean	Convergence	Avg MSE	Avg Coverage
M1	SEFA	SEFA	PSEM	3	33	33.4	100%	0.0033	92.9%
M2	Target-alf	SEFA	PSEM	3	33	33.1	100%	0.0036	94.7%
M3	DSEFA	DSEFA	PSEM	4	24	24.1	100%	0.0039	94.6%
M4	Target-alf	Bifactor	PSEM	4	24	21.9	85%	0.0131	91.0%
M5	Target	Bifactor	ESEM	4	24	22.5	60%	0.0134	86.7%
M6	Geomin	EFA	PSEM	4	24	23.5	88%	0.0164	84.2%
M7	Bi-geomin	Bifactor	PSEM	4	24	21.8	81%	0.0191	86.0%
M5	Target	Bifactor	ESEM	3	33	33.8	100%	0.0222	34.1%
M8	Bi-geomin	Bifactor	ESEM	4	24	22.5	60%	0.0256	52.1%

5.1 Simulation study with SEFA data generation

Figure 7 shows how the data are generated for the first simulation study. The figure also includes the estimation of the SEFA model as well as the computation of the general factor loadings. These represent the indirect effects of the second-order factor on the indicators and are computed as additional parameters. Exact values are computed and set as starting values for the purpose of obtaining proper MSE (mean squared error) and coverage in the Mplus simulation study. These are not necessary in a real-data analysis. The indirect effects of the second-order factor can alternatively be obtained with the MODEL INDIRECT command in Mplus. We compare 9 different estimation methods. The results are presented in Table 1, ordered from best to worst in terms of MSE. The table includes the rotation used for the EFA: Geomin, Bi-geomin, Target, Target-ALF, SEFA and DSEFA. Also included is the model type: bi-factor EFA, SEFA, DSEFA, or EFA. All bi-factor and EFA models are estimated as orthogonal. The target rotation specifies all non-major loadings as targets, including the one small cross-loading, for a total of 24 targets. The Geomin rotation has one additional parameter: the smoothing value in the Geomin function. This is set by default in Mplus to 0.01, which offers a balance between bias in the point estimates and bias in the standard errors. Because this simulation study focuses on point estimates and MSE, each Geomin-based simulation study was estimated with this default setting and also with a smoothing parameter of 0.0001, which typically yields better point estimates. In most cases, the differences were not large, but in some they were substantial. Table 1 contains the better of the two results in terms of MSE. Specifically, for the bi-geomin rotation, a smoothing value of 0.5 has been recommended to minimize factor correlations (Morin, 2023). When applying this alternative smoothing value, the bi-geomin rotation showed improved performance in the first simulation but substantially poorer performance in the second. Given these inconsistent results, we conclude that this smoothing parameter is not suitable for the purposes of the present comparative simulation study.

Since we generate the data as a 3-factor SEFA, equation (7) states that the model is also a 4-factor bi-factor model. The one small cross-loading in the SEFA first-order loading matrix Λ_1 translates to one small cross-loading in the specific factors loading matrix, even though the first-order loading matrix is not on the same scale as the specific factors loading matrix due to differences in factor variances. The variances of the residuals in the SEFA model (< 1) are not the same as the variances of the specific factors ($= 1$). For the general factor, however, the scale is the same. Thus, we evaluate the quality of model estimation via the factor loadings of the general factor. In the SEFA model, these are computed as the indirect loadings from the second-order factor, as illustrated in Figure 7. For the DSEFA model, we use equation (17). The last two columns of Table 1 are based on the general factor loadings. The estimators' performance with respect to the specific factor loadings is comparable to their performance for the general factor loadings.

Generally, we expect that a bi-factor EFA with 4 factors should be able to recover the true generating parameter values, even though a 3-factor EFA fit is also acceptable. We have also included a 3-factor bi-factor model in this simulation study. For this model, the targets are placed on the first two factors while the third specific factor is removed.

We have also included the number of factors in Table 1. This is the number of factors used for the unrotated EFA and determines the degrees of freedom and the chi-square test of fit. As shown in Table 1, the average chi-square value is not identical across the 9 methods. Theoretically, that column should contain just two values: one for the 3-factor EFA and one for the 4-factor EFA. This is not the case here because of different convergence rates and because of small deviations associated with the weight of the penalty for PSEM models. In all cases, the average test of fit matches the degrees of freedom closely, as expected.

The framework and convergence rates are also included in Table 1. In principle, given the same model and rotation, it should not matter which framework is used. However, there is a key difference between the two frameworks: PSEM and ESEM. In ESEM, model estimation is first conducted for the unrotated EFA, and because of this, starting values given in the input are essentially ignored. In PSEM, the unrotated model is not estimated. The final model is estimated directly, and starting values given in the input are utilized. The unrotated model is generally more difficult to estimate than the direct PSEM model.

The results in Table 1 are striking. Method M8, Bi-geomin Bifactor model with ESEM, which is currently the most widely used approach, performs the worst. Switching the framework to PSEM (method M7) somewhat improves the results. Changing the rotation to Target (method M5) improves the situation even more. Note that M5 uses more information than just the number of factors as in M8, namely the location of the targets. Changing the rotation to Target-ALF (method M4) further improves the results and appears to be the best method for a direct Bifactor model. Reducing the number of factors to 3 (using method M5) does not improve the results because one of the specific factors is combined with the general factor. The two SEFA methods (M1 and M2) perform the best. This is likely due to the fact that these models provide more parsimonious fits to the data and the correct number of factors. The simulation study setup for these two methods are given in Figures 7 and 8. The DSEFA method (M3) yields nearly as good performance, as expected, and is the best method for estimating the 4-factor bi-factor EFA model through reparameterization. Comparing methods M6, which is identical to TYPE=EFA in Mplus, and M7—that is, replacing the Bi-geomin rotation with the generic Geomin rotation—also improves the results. This further exposes the weakness of the Bi-geomin rotation. Its design to exclude the general factor from the rotation is flawed for this model and this part of the parameter space. The PSEM framework appears universally more robust in terms of yielding higher convergence rates.

We presented this simulation study as based on 3-factor SEFA data generation; however, the data can also be generated from the equivalent (7) 4-factor bi-factor model. If we perturbed the specific factor loadings slightly, the 3-factor EFA would not hold and a 4-factor EFA would be the true model. If the perturbation is small, the quality of the results for all of the above methods will not change substantially. Thus, Table 1 applies not just to 3-factor SEFA data generation, but also to nearby 4-factor bi-factor data-generating models. We conclude that if a bi-factor model is close to a SEFA model, we can generally expect that bi-factor EFA will have poor replicability and that SEFA/DSEFA models should be preferred over bi-factor EFA.

Figure 7: Simulation study for 3-factor SEFA model with Geomin rotation: method M1

```
MONTECARLO:
  NAMES=y1-y12; NOBS=500;
  NREPS=100;

ANALYSIS: ESTIMATOR=MLR; rotation=sefa;

MODEL MONTECARLO:
  f1 BY y1-y4*.7 y5-y12*0;
  f2 BY y1*0.25 y2-y4*0 y5-y8*.7 y9-y12*0;
  f3 BY y1-y8*0 y9-y12*.7;
  y1-y12*.4;
  f by f1*0.5 f2*0.6 f3*0.7;
  f1*0.75 f2*0.64 f3*0.51; f@1;

MODEL:
  f1-f3 by y1-y12(*1);
  f1 BY y1-y4*.7 y5-y12*0(a1-a12);
  f2 BY y1*0.25 y2-y4*0 y5-y8*.7 y9-y12*0(b1-b12);
  f3 BY y1-y8*0 y9-y12*.7(c1-c12);
  y1-y12*.4;
  f by f1*0.5 f2*0.6 f3*0.7 (L1-L3);
  f1*0.75 f2*0.64 f3*0.51; f@1;

model constraint:
  ! Total indirect effect from F to each Y indicator
  NEW(IE1*.5 IE2-IE4*.35 IE5-IE8*.42 IE9-IE12*.49);
  do(#,1,12) IE#=L1*a#+L2*b#+L3*c#;
```

Figure 8: Simulation study for 3-factor SEFA model with Target-alf rotation: method M2

```
MONTECARLO:
  NAMES=y1-y12; NOBS=500;
  NREPS=100;

ANALYSIS: ESTIMATOR=MLR;

MODEL MONTECARLO:
  f1 BY y1-y4*.7 y5-y12*0;
  f2 BY y1*0.25 y2-y4*0 y5-y8*.7 y9-y12*0;
  f3 BY y1-y8*0 y9-y12*.7;
  y1-y12*.4;
  f by f1*0.5 f2*0.6 f3*0.7;
  f1*0.75 f2*0.64 f3*0.51; f@1;

MODEL:
  f1 BY y1-y4*.7 y5-y12*0(a1-a12);
  f2 BY y1*0.25 y2-y4*0 y5-y8*.7 y9-y12*0(b1-b12);
  f3 BY y1-y8*0 y9-y12*.7(c1-c12);
  y1-y12*.4;
  f by f1*0.5 f2*0.6 f3*0.7 (L1-L3);
  f1*0.75 f2*0.64 f3*0.51 (v1-v3); f@1;

model prior:
  a5-a12~ALF(0,1);
  b1-b4~ALF(0,1);
  b9-b12~ALF(0,1);
  c1-c8~ALF(0,1);

model constraint:
  do(#,1,3) v#=1-L#*L#;

  ! Total indirect effect from F to each Y indicator
  NEW(IE1*.5 IE2-IE4*.35 IE5-IE8*.42 IE9-IE12*.49);
  do(#,1,12) IE#=L1*a#+L2*b#+L3*c#;
```

5.2 Simulation study with bi-factor EFA data generation

In the second simulation, we use a slightly larger perturbation for the specific loadings. Instead of using equal loadings, we use loadings that are all different and 0.3 units apart. Figure 9 shows the data generation as well as the setup for a bi-factor EFA estimation with the DSEFA reparameterization. The transformation (reparameterization) of the DSEFA model into the bi-factor model is provided in the MODEL CONSTRAINT command for illustration purposes only. In real data analysis, Mplus will automatically compute the parameter estimates for the two equivalent models.

The results from several other estimation methods are given in Table 2. Method M1, 3-factor SEFA, is included to show that this model does not collapse to a 3-factor SEFA—that is, it is not too close to a SEFA model. With an average chi-square value of 59.5 and 33 degrees of freedom, the 3-factor model is rejected 80% of the time. In this situation, the winning method is the DSEFA estimation method. This was the best Bifactor method in the previous simulation study as well. The bi-factor model with the Target-ALF rotation is second best, as it was in the previous simulation study. This method is illustrated in Figure 10. The Target rotation is the third best bi-factor method and again the Bi-geomin rotation is the worst. Interestingly, the SEFA method with 4 factors (method M1) performs fairly well. This occurs even though the SEFA method uses 4 specific factors rather than 3 as in the data generation. The additional specific factor is primarily compensating for the non-proportionality in the loadings and would be difficult to interpret in practical settings, with all of its loadings being small and barely significant. Nevertheless, the SEFA method here is nearly as effective in identifying the general factor. We interpret this as an advantage of the SEFA method, demonstrating a level of robustness.

In conclusion, the above simulations demonstrate that the DSEFA/SEFA models are serious contenders when considering EFA with a general factor. The models are robust, easy to estimate, and provide accurate parameter estimates. Direct bi-

Table 2: EFA methods (Sorted by MSE Performance), Data generated from a 4-factor Bifactor

Method	Rotation	Model	Framework	Factors	DF	χ^2 Mean	Conver-gence	Avg MSE	Avg Coverage
M3	DSEFA	DSEFA	PSEM	4	24	25.2	100%	0.0069	93.8%
M4	Target-alf	Bifactor	PSEM	4	24	25.1	99%	0.0152	97.1%
M1	SEFA	SEFA	PSEM	4	24	25.4	98%	0.0167	93.4%
M5	Target	Bifactor	ESEM	4	24	25.6	97%	0.0255	87.2%
M8	Bi-geomin	Bifactor	ESEM	4	24	25.6	97%	0.0270	93.5%
M1	SEFA	SEFA	PSEM	3	33	59.5	97%	0.0941	86.3%

factor EFA models (excluding the DSEFA reparameterization), on the other hand, may in some circumstances yield quite inaccurate parameter estimates. Unless loading non-proportionality is clearly visible, as in the case where some indicators load only on the general factor, a real-data-based simulation study should be performed to ensure replicability of estimates. It is also now abundantly clear that the rotations for these models, in order of preference, should be: Target-ALF, Target, and least desirable, Bi-Geomin. Finally, we have illustrated here that a large number of alternative options now exist for general factor EFA estimation, and many of the new options are much better than those used traditionally.

Figure 9: Simulation study for 4-factor (3 specific and 1 general) bi-factor model with DSEFA reparameterization: method M3

```
MONTECARLO:
  NAMES=y1-y12; NOBS=500; NREPS=100;

ANALYSIS: ESTIMATOR=MLR; rotation=dsefa;

! Generating data from a bi-factor model
MODEL MONTECARLO:
  f1 BY y1*.6 y2*.9 y3*1.2 y4*1.5;
  f2 BY y1*.2 y5*.5 y6*.8 y7*1.1 y8*1.4;
  f3 BY y9*.5 y10*.8 y11*1.1 y12*1.4;
  f by y1*0.5 y2-y4*0.35 y5-y8*0.42 y9-y12*0.49;
  y1-y12*.4 f1-f3@1 f@1;

! Estimating DSEFA, labels needed for reparam
MODEL:
  f f1-f3 by y1-y12(*1);
  f1 BY y1-y12(a1-a12);
  f2 BY y1-y12(b1-b12);
  f3 BY y1-y12(c1-c12);
  f by y1-y12(d1-d12);
  f by f1-f3 (L1-L3);

model constraint:
  ! Reparameterization from DSEFA to bi-factor
  ! Starting values match model montecarlo to obtain MSE
  ! Total effect from general to each indicator
  new(TE1*.5 TE2-TE4*.35 TE5-TE8*.42 TE9-TE12*.49);
  do(#,1,12) TE#=L1*a#+L2*b#+L3*c#+d#;
  !stand loadings for first specific
  new(p1*.6 p2*.9 p3*1.2 p4*1.5 p5-p12*0);
  do(#,1,12) p#=a#*sqrt(1-L1*L1);
  !stand loadings for second specific
  new(q1*.2 q2-q4*0 q5*.5 q6*.8 q7*1.1 q8*1.4 q9-q12*0);
  do(#,1,12) q#=b#*sqrt(1-L2*L2);
  !stand loadings for third specific
  new(r1-r8*0 r9*.5 r10*.8 r11*1.1 r12*1.4);
  do(#,1,12) r#=c#*sqrt(1-L3*L3);
```

Figure 10: Simulation study for 4-factor (3 specific and 1 general) bi-factor model with Target-ALF rotation: method M4

```
MONTECARLO:
  NAMES=y1-y12; NOBS=500;
  NREPS=100;
MODEL MONTECARLO:
  f1 BY y1*.6 y2*.9 y3*1.2 y4*1.5 y5-y12*0;
  f2 BY y1*0.2 y2-y4*0 y5*.5 y6*.8 y7*1.1 y8*1.4 y9-y12*0;
  f3 BY y1-y8*0 y9*.5 y10*.8 y11*1.1 y12*1.4;
  y1-y12*.4;
  f by y1*0.5 y2-y4*0.35 y5-y8*0.42 y9-y12*0.49;
  f1-f3@1 f@1;
MODEL:
  f1 BY y1*.6 y2*.9 y3*1.2 y4*1.5 y5-y12*0(p1-p12);
  f2 BY y1*0.2 y2-y4*0 y5*.5 y6*.8 y7*1.1 y8*1.4 y9-y12*0(p13-p24);
  f3 BY y1-y8*0 y9*.5 y10*.8 y11*1.1 y12*1.4(p25-p36);
  y1-y12*.4;
  f by y1*0.5 y2-y4*0.35 y5-y8*0.42 y9-y12*0.49;
  f f1-f3 with f1-f3@0;
  f@1 f1-f3@1;
model prior:
  p5-p16~ALF(0,1);
  p21-p32~ALF(0,1);
```

6 An empirical example

We illustrate the SEFA and DSEFA methods with data from Holzinger and Swineford (1939). Test scores on 24 different mental ability measures were obtained from a total of 301 7th- and 8th-grade students in two schools. Muthén and Asparouhov (2012) use 19 of the items that are hypothesized to measure four domains: spatial abilities, verbal abilities, speed, and memory. Here, we will analyze this set of 19 items as well as the full set of 24 items separately. Holzinger and Swineford (1939) hypothesized the existence of a general factor affecting all test scores, and therefore SEFA and DSEFA are suitable modeling approaches.

6.1 SEFA for 19 items

First we analyze the 19-item set of indicators. Figure 11 shows the input file for estimating the SEFA model. All variables are standardized in the DEFINE command to avoid suboptimal rotation due to variables being on substantially different scales. Figure 12 contains SEFA results and Figure 13 contains the results of the standard oblique EFA for comparison. The loading structure between the two models is largely unchanged. EFA and SEFA 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 SEFA is suitable for the Holzinger–Swineford data and possibly provides a more practical interpretation of a general ability factor and specific domain factors. EFA and SEFA 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 14 contains the results for the bi-factor EFA with 4 factors and Figure 15 contains the results for 5 factors both estimated with the bi-geomin rotation. The 4 factor model in Figure 14 is equivalent to the Figure 12 and Figure 13 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.

6.2 DSEFA for 19 items

We can also use the DSEFA model to estimate a bi-factor EFA with 4 specific factors and 1 general factor. Fit statistics for the 19-item example are given in Table 3. Note that fit statistics do not depend on the rotation method but only on the number of EFA factors. In SEFA models, the number of EFA factors is the number of first order factors while in DSEFA models the number of EFA factors is the number of first order factors plus the number of second order factors. That is, the fit of SEFA with 4 first order factors matches the fit of EFA with 4 factors, while the fit of DSEFA with 4 first order factors matches the fit of EFA with 5 factors.

In this example, BIC favors the 4-factor EFA over the 5-factor EFA. The likelihood ratio test between the 4-factor EFA and the 5-factor EFA does not reject the 4-factor solution. This suggests that the bi-factor model with 4 specific factors and one general is likely to collapse to the 4-factor SEFA model. As our simulation studies indicate, DSEFA is particularly well suited to this situation. When the bi-factor model degenerates toward a SEFA solution, DSEFA can be expected to recover the correct specific factor structure more accurately than direct bi-geomin rotation. Indeed, no significant direct effects are found in DSEFA, confirming that the bi-factor EFA with 4 specific factors and one general collapses to the 4-factor SEFA model. In addition, the DSEFA bi-factor EFA reparameterization, presented in Figure 16, shows that all four specific factors are properly recovered by DSEFA, outperforming the bi-geomin rotation.

Figure 11: SEFA input file for Holzinger and Swineford 19-item example

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

variable: names are 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: rotation=sefa;

model:
    spatial verbal speed memory by visual-figurew(*1);
    f0 by spatial-memory;
```

Figure 12: SEFA results for Holzinger and Swineford 19 items with 4 first order factors

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*

Table 3: Fit statistics for Holzinger and Swineford 19-item example

EFA factors	Log-Likelihood	Parameters	BIC	Chi-Square	p-value
4	-7106.3	108	14829	131.9	.021
5	-7091.3	123	14885	112.7	.028

Figure 13: EFA results for Holzinger and Swineford 19 items with 4 factors

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 14: Bi-factor EFA results for Holzinger and Swineford 19 items 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 15: Bi-factor EFA results for Holzinger and Swineford 19 items 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

Figure 16: Bi-factor EFA results for Holzinger and Swineford 19 items with 5 factors via DSEFA

ROTATED LOADINGS (* significant at 5% level)

	General	Spatial	Verbal	Speed	Memory
VISUAL	0.433*	0.527*	0.185*	0.013	0.062
CUBES	0.268*	0.396*	0.062	−0.119	−0.023
PAPER	0.223*	0.419*	0.151*	0.012	−0.034
FLAGS	0.285*	0.585*	0.005	0.027	0.130
GENERAL	0.345*	0.030	0.766*	0.082	−0.027
PARAGRAP	0.429*	0.015	0.695*	0.025	0.087
SENTENCE	0.507*	−0.117	0.727*	−0.018	−0.067
WORDC	0.477*	0.030	0.573*	0.014	0.039
WORDM	0.375*	0.109	0.759*	0.014	0.073
ADDITION	0.120	−0.024	0.088	0.772*	0.130
CODE	0.507*	0.032	0.105	0.454*	0.139
COUNTING	0.283*	0.179	0.012	0.584*	−0.024
STRAIGHT	0.586*	0.261	−0.062	0.385*	−0.080
WORDR	0.336	−0.061	0.018	−0.044	0.581*
NUMBERR	0.117	0.157	−0.041	0.024	0.577*
FIGURER	0.416	0.259	0.039	−0.002	0.404*
OBJECT	0.221*	−0.027	−0.008	0.311*	0.509*
NUMBERF	0.222	0.167	0.064	0.182	0.404*
FIGUREW	0.354*	0.044	0.113	0.022	0.262

Table 4: Fit statistics for Holzinger and Swineford 24-item example

EFA factors	Log-Likelihood	Parameters	BIC	Chi-Square	p-value
4	-8893.5	138	18575	283.3	.0000
5	-8862.5	158	18627	238.6	.0002

6.3 24-item analysis

Next, we analyze all 24 items. The fit statistics are given in Table 4. The 4-factor EFA continues to yield the lowest BIC, but a likelihood ratio test against the 5-factor EFA now rejects the 4-factor solution. Furthermore, the DSEFA analysis reveals significant direct effects from the second-order factor to three of the five additional indicators, suggesting that the 5-factor DSEFA/bi-factor model does not fully collapse to the 4-factor SEFA model in this larger item set. The 5 extra items mostly load on the general factor which likely prevents DSEFA from collapsing to SEFA. Figure 17 shows the input file for estimating DSEFA for the 24-item model. Figures 18–20 present the results of the 24-item analysis under three approaches: bi-factor EFA via DSEFA, bi-factor EFA via bi-Geomin(orthogonal), and the SEFA model.

Compared to the 19-item analysis, the three methods show somewhat better agreement in the 24-item case, consistent with the pattern observed in our simulation studies when collapsing is not present. Nevertheless, the DSEFA solution provides the clearest rotation of the general factor and the four specific factors, matching the psychometric theory as well as our simulation study findings. Since DSEFA and bi-geomin rotations are both orthogonal here, the sum of the squared loadings in each row for both Figures 18 and 19 are identical and equal R^2 of the 5 factor EFA for each indicator. We can also compute the sum of squared loadings just for the specific factors. The average of these across indicators is 0.29 for DSEFA and 0.23 for bi-geomin. This represents a 25% increase in DSEFA's ability to more clearly orient and measure the specific factors.

Figure 17: DSEFA input file for Holzinger and Swineford 24-item example

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

variable: names are 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-arithmet;

define: standardize visual-arithmet;

analysis: rotation=dsefa;

model: gen spatial verbal speed memory by visual-arithmet(*1);
```

Figure 18: Bi-factor EFA results for Holzinger and Swineford 24 items with 5 factors via DSEFA

ROTATED LOADINGS (* significant at 5% level)

	General	Spatial	Verbal	Speed	Memory
VISUAL	0.414*	0.532*	0.182*	0.109	0.046
CUBES	0.274*	0.425*	0.044	−0.064	−0.047
PAPER	0.251*	0.371*	0.112	0.075	−0.069
FLAGS	0.367*	0.535*	−0.056	0.086	0.070
GENERAL	0.398*	−0.009	0.728*	0.040	−0.068
PARAGRAP	0.441*	0.026	0.690*	0.005	0.062
SENTENCE	0.360*	0.006	0.809*	0.012	−0.025
WORDC	0.427*	0.077	0.602*	0.022	0.039
WORDM	0.474*	0.068	0.696*	−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*	0.040	−0.023	0.540*	−0.062
STRAIGHT	0.347*	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*	0.018	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	0.520*	0.266*	0.155	−0.137*	0.102
NUMERIC	0.664*	0.111	0.049	0.140	−0.025
PROBLEMR	0.529*	0.233	0.321*	−0.057	0.018
SERIES	0.674*	0.278*	0.179*	−0.021	−0.001
ARITHMET	0.689*	−0.121	0.172	0.088	0.055

Figure 19: Bi-factor EFA results for Holzinger and Swineford 24 items with 5 factors via bi-Geomin

ROTATED LOADINGS (* significant at 5% level)

	General	Spatial	Verbal	Speed	Memory
VISUAL	0.616*	0.361*	0.003	0.016	−0.006
CUBES	0.415*	0.239*	−0.095	−0.140	−0.096
PAPER	0.387*	0.259	−0.005	0.011	−0.108
FLAGS	0.534*	0.313	−0.233*	−0.002	0.013
GENERAL	0.502*	−0.035	0.661*	0.019	−0.083
PARAGRAP	0.552*	−0.024	0.609*	−0.019	0.046
SENTENCE	0.489*	0.012	0.739*	−0.017	−0.037
WORDC	0.541*	0.011	0.512*	−0.009	0.018
WORDM	0.595*	−0.018	0.598*	−0.072	−0.011
ADDITION	0.290*	−0.349	−0.015	0.677*	0.011
CODE	0.432*	0.038	0.149*	0.507*	0.169*
COUNTING	0.379*	−0.026	−0.070	0.535*	−0.083
STRAIGHT	0.471*	0.254*	−0.030	0.464*	−0.029
WORDR	0.301*	−0.071	0.027	−0.022	0.588*
NUMBERR	0.265*	0.004	−0.167*	−0.007	0.496*
FIGURER	0.541*	0.054	−0.102	−0.023	0.332*
OBJECT	0.256*	−0.073	−0.027	0.296*	0.493*
NUMBERF	0.333*	0.027	−0.022	0.166*	0.363*
FIGUREW	0.423*	−0.094	0.039	−0.019	0.221
DEDUCT	0.605*	0.021	0.025	−0.179*	0.064
NUMERIC	0.661*	−0.142	−0.049	0.138	−0.057
PROBLEMR	0.631*	0.032	0.193*	−0.101	−0.018
SERIES	0.751*	−0.009	0.032	−0.059	−0.046
ARITHMET	0.620*	−0.327	0.127	0.128	0.046

Figure 20: SEFA results for Holzinger and Swineford 24 items with 4 first order factors

ROTATED LOADINGS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
VISUAL	0.625*	0.107	0.019	−0.012
CUBES	0.569*	−0.004	−0.120	−0.073
PAPER	0.454*	0.061	0.014	−0.119
FLAGS	0.685*	−0.158*	0.026	0.032
GENERAL	−0.023	0.850*	0.049	−0.104*
PARAGRAP	0.019	0.800*	0.009	0.042
SENTENCE	−0.047	0.917*	−0.020	−0.073
WORDC	0.092	0.687*	0.023	0.014
WORDM	0.097	0.813*	−0.029	−0.008
ADDITION	−0.250*	0.009	0.852*	0.042
CODE	0.003	0.172*	0.537*	0.134
COUNTING	0.109	−0.046	0.650*	−0.090
STRAIGHT	0.316*	−0.010	0.468*	−0.061
WORDR	−0.030	0.069	−0.012	0.662*
NUMBERR	0.118	−0.148*	0.015	0.561*
FIGURER	0.365*	−0.001	0.035	0.411*
OBJECT	−0.109	−0.039	0.326*	0.507*
NUMBERF	0.083	0.003	0.192*	0.377*
FIGUREW	0.150	0.150*	0.060	0.290*
DEDUCT	0.461*	0.199*	−0.075	0.142*
NUMERIC	0.362*	0.126	0.317*	0.043
PROBLEMR	0.392*	0.381*	−0.014	0.031
SERIES	0.523*	0.246*	0.086	0.045
ARITHMET	0.099	0.325*	0.322*	0.155

F0 BY

Spatial	0.734*
Verbal	0.577*
Speed	0.511*
Memory	0.437*

FACTOR CORRELATIONS (* significant at 5% level)

	Spatial	Verbal	Speed	Memory
Spatial	1.000			
Verbal	0.424*	1.000		
Speed	0.375*	0.295*	1.000	
Memory	0.321*	0.252*	0.223*	1.000
F0	0.734*	0.577*	0.511*	0.437*

6.4 Model Selection and Estimation Procedure

The empirical example above demonstrates that model preference is context-dependent: the SEFA model proves superior in the 19-item analysis, while the DSEFA model is preferable in the 24-item analysis. Notably, both models support interpretation as either a second-order or bi-factor model with a general ability factor influencing all observed indicators. In this section, we outline a general workflow for estimating EFA models with a general latent dimension.

The first step is to determine the number of EFA factors needed to fit the data well. This topic has been extensively covered in the literature, and various methods have been proposed. A recent summary and accompanying simulation studies are reported in Asparouhov and Muthén (2024b), which outline five reliable methods; two of these—BIC and the sequential likelihood ratio test—were employed in the example above. Such methods typically yield one or two consecutive candidate values for the number of factors m. When two values are in contention, i.e., m and $m+1$, the remaining steps should be conducted for both. At this stage, the Mplus default Geomin rotation can be used within either the ESEM or EFA framework.

The second step is to consider adding residual covariances between observed indicators to the ESEM model. Model fit, modification indices and residual output available in Mplus—which compares estimated and observed variance-covariance matrices for the indicators—can assist in this process. However, the PSEM framework offers a more elegant and reliable solution. The procedure, outlined in Section 4.5 of Asparouhov and Muthén (2024a), aligns residual covariances toward zero while simultaneously estimating the EFA model. At this stage, attention should also be given to factors loading on fewer than three indicators, as it may be possible to replace such factors with a single residual covariance parameter. Likelihood ratio tests can be used to compare such alternative models.

The third step is to estimate the EFA using SEFA and DSEFA rotations. When available, substantive theory should serve as the primary guide for model selection. Depending on the clarity of the factor number decision, there will be either two or four models to evaluate. In the absence of substantive theory, the first priority should be to narrow the candidate set on technical grounds. Relevant considerations include convergence failures or Mplus warnings such as Heywood cases or a poorly conditioned information matrix; EFA factors with very few or no significant or substantial loadings; and poor interpretability, for example when items load substantially on many or all factors. SEFA and DSEFA models should have at least two significant second-order loadings to be considered meaningful. Models with unusually large standard errors, indicative of near non-identification, should also be discarded. Sensitivity analyses for the penalty weights (prior variances) are additionally recommended.

We now describe the comparisons that can be made among the remaining models. First, DSEFA(m) with m first-order factors is nested above SEFA(m) with m first-order factors, since DSEFA simply adds direct effects from the second-order factor to the indicators. Accordingly, if any direct effect is significant, preference should be given to the DSEFA model; if none are significant, the SEFA model is preferred. Note that DSEFA(m) and SEFA(m) do not have identical data fit, as DSEFA(m) fits equivalently to EFA($m + 1$). When the difference in fit is substantial, it should be factored into the model selection decision.

Next, consider the comparison between DSEFA(m) and SEFA($m+1$). These two models have identical data fit, and SEFA($m+1$) can be viewed as nesting DSEFA(m): fixing a second-order loading in SEFA($m + 1$) to 1 renders it equivalent to DSEFA(m). Therefore, if all second-order loadings are statistically different from 1—equivalently, if all first-order factor residual variances are statistically significant—preference should be given to SEFA($m + 1$). Conversely, if a second-order loading is not statistically different from 1—equivalently, if a first-order factor residual variance is not statistically significant—preference should be given to DSEFA(m).

In complex examples involving many factors and indicators, manually evaluating the first-order loading matrix may be impractical. The Geomin simplicity function provides a useful summary of overall model simplicity. The lower the simplicity value, the easier the interpretation is. Tables 5 and 6 report Geomin simplicity values for EFA, SEFA, and DSEFA models for the 19-item and 24-item examples. Within each row, models have identical data fit, making their simplicity values directly comparable. We also report the percentage increase relative to the minimum value in each row; differences smaller than 5% should be considered negligible, as they are unlikely to be of substantive importance. In principle, formal statistical testing based on confidence intervals for each simplicity value could also be used to assess significance, though this will rarely be necessary in practice. The substantively motivated model selections are bolded in each table; in both examples, the selected model achieves a simplicity value at or near the row minimum, confirming the appropriateness of those choices.

Tables 5 and 6 also show that EFA and SEFA yield nearly identical simplicity values, which can be interpreted as evidence that the EFA factor correlation structure is compatible with a second-order factor model. Theoretically, EFA simplicity should always be no greater than SEFA simplicity; the small reversals observed in both tables likely reflect SEFA's ability to trade a marginal reduction in data fit for a gain in simplicity, and such differences should be disregarded. This provides an additional justification for ignoring simplicity differences smaller than 5%. Note that simplicity values across rows are not comparable. Furthermore, a model such as DSEFA can be parameterized either as a bi-factor model or as a second-order factor model, and the corresponding loading matrices yield different simplicity values; it is therefore important not to compare the simplicity of a bi-factor parameterization with that of a second-order parameterization. We do not recommend comparing DSEFA bi-factor simplicity with that obtained from a bi-Geomin rotation, as bi-Geomin rotation is driven entirely by minimizing the bi-Geomin criterion, whereas DSEFA incorporates an additional stabilizing component that penalizes direct effects. Our simulations indicate that relying solely on bi-Geomin minimization tends to skew specific factor loadings toward lower values, yielding inferior results.

Table 5: Geomin simplicity (increase over lowest value in the row) Holzinger and Swineford 19-item example

EFA factors m	EFA(m)	SEFA(m)	DSEFA($m-1$)
4	.708 (1%)	**.701**	.770 (10%)
5	.625 (0%)	.622	.653 (5%)

Table 6: Geomin simplicity (increase over lowest value in the row) Holzinger and Swineford 24-item example

EFA factors m	EFA(m)	SEFA(m)	DSEFA($m-1$)
4	.949 (1%)	.940	1.048 (11%)
5	.804	.804	**.813 (1%)**

7 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. For this illustration we use the 19-item SEFA model but similar analysis can be conducted with the 24-item DSEFA model.

We consider a sequence of four models. Model M1 is the ESEM model where all factors are regressed on all covariates. Model M2 is SEFA 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 SEFA as well as the power and complexities of the PSEM methodology.

Model M3 is SEFA 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, 2024a). 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

$$Y = \nu + \Lambda_1 F + \varepsilon, \tag{21}$$

$$F = B_1 X + \Lambda_2 \eta + \xi, \tag{22}$$

$$\eta = B_2 X + \zeta, \tag{23}$$

$$\varepsilon \sim N(0, \theta), \xi_i \sim N(0, 1 - \lambda_{2i}^2), \zeta \sim N(0, 1), \tag{24}$$

$$\Lambda_1 \sim \text{Geomin}(4, 0.1), B_1 \sim \text{ALF}(0, 1) \tag{25}$$

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

Figure 21 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 22 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 SEFA offers a superior summary of the covariate effects on the factors when compared to the general ESEM/EFA analysis with factor specific effects. 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 12 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 SEFA 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 SEFA 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).

Figure 21: SEFA with covariates

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


**** Model M2 ****
  analysis: rotation=sefa;

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


**** Model M3 ****
  analysis: rotation=sefa;

  model:
      spatial verbal speed memory by visual-figurew(*1);
      f by spatial verbal speed memory;
      f on female grade school;
      spatial verbal speed memory on female grade school (b1-b12);

  model priors: b1-b12~ALF(0,1);


**** Model M4 ****
  analysis: rotation=sefa;

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

Figure 22: SEFA with covariates results

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

7.1 Degrees of Freedom and Number of Free Parameters in PSEM Models

The degrees of freedom in PSEM models is a nuanced concept, estimated numerically (see Asparouhov and Muthén, 2024a). While straightforward in many cases, it can be difficult to interpret in others. The degrees of freedom plus the number of free parameters is as usual the number of parameters in the unrestricted model ($p(p+1)/2+p$ with p dependent variables which includes the means). The number of free parameters, however, is not the number of parameters in the model. In the presence of a penalty/rotation function, in the PSEM framework, minimizing the penalty is equivalent to a number of parameter constraints (first derivatives=0), i.e., some parameters are not independent free parameters but functions of other model parameters. These constraints are not all independent so it is not possible to just count the number of penalty derivatives, rather the rank of the constraint matrix becomes the difference between the number of model parameters and the number of free parameters. This rank is essentially numerically estimated.

Consider the oblique EFA and SEFA models for the 19-item analysis, both of which have 108 free parameters. The EFA model begins with 120 parameters, and the oblique rotation eliminates/identifies 12 — leaving 108 free. The SEFA model begins with 122 parameters, and its rotation (which consists of penalty and model constraints) eliminates 14. Thus, different models are not always affected equally by rotation. This is already apparent when comparing oblique and orthogonal EFA: in the oblique case, the rotation additionally identifies all factor correlations (6 parameters), whereas the orthogonal EFA starts with 114 parameters and the rotation identifies 6, again yielding 108 free parameters.

When EFA models are embedded within structural models — as in the PSEM framework — the counts of free parameters and degrees of freedom become more complex. This complexity does not arise in the ESEM framework in Mplus, which is more restrictive and limited to models that can be fully rotated without altering the model's structure. For example, ESEM does not permit regressing a single exploratory factor on one covariate (as in model M4); all exploratory factors must be regressed on that covariate simultaneously (as in model M1).

The more flexible PSEM framework introduces the concept of *native model rotation* (discussed in Section 6.7 of Asparouhov and Muthén, 2024a), whereby structural restrictions can serve as rotation criteria. In some models, the structural constraints alone are sufficient to fully identify the model, rendering an explicit rotation criterion unnecessary. In others, the constraints only partially identify the model — acting as a *partial native rotation* — thereby reducing, but not eliminating, the identifying burden placed on the rotation criterion.

The four models in Figure 21 illustrate this. Models M1 and M3 each have 120 free parameters, reflecting 12 additional regression coefficients. Model M2, however, has 114 free parameters — six more than the baseline SEFA model — despite adding only three new regression parameters. This is because the MIMIC structure of the second-order model acts as a partial native rotation, reducing the identifying role of the first-order loadings rotation. More generally, a SEFA M2 model with m first-order factors and q covariates predicting the second-order factor adds $m+q-1$ free parameters beyond the SEFA model without covariates: the first covariate contributes m new parameters, while each additional covariate contributes one. Model M4, in turn, has 117 free parameters — three more than M2 — despite adding only one direct effect.

These examples highlight a counterintuitive property: adding a single parameter to a model may count as several free parameters, or conversely, as none at all. For instance, adding a direct effect from gender to verbal in M4 counts as one additional free parameter, but subsequently adding a direct effect from grade to verbal counts as zero — meaning it would not improve model fit, though it may influence the first-order rotation. When a parameter is added to a model, the structural constraints in the model are altered and if these changes interact with the rotation, the number of free parameters may change by a number different from 1.

Rather than seeking a closed-form account of how loading rotation and native partial rotation jointly determine degrees of freedom, one can simply rely on the numerical procedure — based on the rank of the information matrix — that Mplus automatically computes. When further verification is needed, a Monte Carlo simulation study provides a practical check: generating large data from the PSEM model and reanalyzing it with the same model across a sufficient number of replications should yield an average chi-square value that matches the reported degrees of freedom.

8 Conclusion

As we have gained more experience with the SEFA models, 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 achieve. 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. The DSEFA model described here is special because it offers a natural way to connect SEFA models and bi-factor EFA models. It also appears to provide the most accurate way to estimate bi-factor EFA models indirectly via a simple reparameterization automatically computed in Mplus. The DSEFA model also provides an elegant technical resolution to the rank deficiency of the Schmid and Leiman (1957) bi-factor EFA estimation.

We have promoted here the view that the SEFA model's natural place is in the PSEM framework, and not in the ESEM framework via the EWC estimation of Morin et al. (2016) and Morin and Asparouhov (2018). Nevertheless, currently the SEFA analysis is primarily being conducted with the EWC approach. Clearly, further practical applications are needed to compare the two methods, as are 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

SEFA models.

Mplus 9.1 model setup that we illustrate here, automates much of the SEFA and DSEFA model estimation. Specifying the rotation as SEFA or DSEFA, automatically creates the needed rotation penalty and model parameters constraints. This setup also triggers an elaborate starting values procedure rooted in EFA traditions. Thus, exhaustive random starting value optimizations typically are no longer needed. It appears that the PSEM based rotation is now as good and as easy to use as the gradient projection algorithms of Jennrich (2002), while also allowing us to break through the ESEM mold of restrictions. Currently, SEFA and DSEFA rotations are available for single level models with continuous and categorical indicators, however, the method can be used with any Mplus model by manually specifying the penalty function and parameter constraints.

The focus of this paper was to provide 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 SEFA 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. (2024a). Penalized Structural Equation Models. Structural Equation Modeling: A Multidisciplinary Journal, 31, 429–454.

[4] Asparouhov, T., & Muthén, B. (2024b). Testing the number of factors in EFA and CFA. Mplus Web Note No.25 https://www.statmodel.com/examples/webnotes/webnote%2025.pdf

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

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

[7] Giordano, C., & Waller, N. G. (2020). Recovering bifactor models: A comparison of seven methods. Psychological Methods, 25, 143–156.

[8] Jennrich, R.I. (2002). A simple general method for oblique rotation. Psychometrika, 67, 7-20.

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

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

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

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

[13] 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.

[14] Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich– Bentler analytic rotations. Multivariate Behavioral Research, 51, 698–717.

[15] Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. Intelligence, 61, 120–129.

[16] Morin, A.J.S. (2023). Exploratory structural equation modeling. In R.H. Hoyle (Ed)., Handbook of Structural Equation Modeling, Second Edition. (New York, NY: The Guilford Press), 503-524.

[17] 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

[18] 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.

[19] Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially-specified target matrices: Applications in exploratory and Bayesian confirmatory factor analysis. Multivariate Behavioral Research, 50, 149–161.

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

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

[22] Reise, S. P., Mansolf, M., and Haviland, M. G. (2023). Bifactor measurement models. Handbook of structural equation modeling. ed. R. H. Hoyle. Second Edition (New York, NY: The Guilford Press), 329–348.

[23] Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. Psychometrika, 22, 53–61.

[24] Waller, N. G. (2018). Direct Schmid Leiman transformations and rank deficient loadings matrices. Psychometrika, 83, 858-870.