



A Simple Two-Step Procedure for Fitting Fully Unrestricted Exploratory Factor Analytic Solutions with Correlated Residuals

Pere J. Ferrando, Ana Hernández-Dorado & Urbano Lorenzo-Seva

To cite this article: Pere J. Ferrando, Ana Hernández-Dorado & Urbano Lorenzo-Seva (19 Dec 2023): A Simple Two-Step Procedure for Fitting Fully Unrestricted Exploratory Factor Analytic Solutions with Correlated Residuals, Structural Equation Modeling: A Multidisciplinary Journal, DOI: [10.1080/10705511.2023.2267181](https://doi.org/10.1080/10705511.2023.2267181)

To link to this article: <https://doi.org/10.1080/10705511.2023.2267181>



© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC



Published online: 19 Dec 2023.



Submit your article to this journal [↗](#)



Article views: 243



View related articles [↗](#)



View Crossmark data [↗](#)

A Simple Two-Step Procedure for Fitting Fully Unrestricted Exploratory Factor Analytic Solutions with Correlated Residuals

Pere J. Ferrando , Ana Hernández-Dorado  and Urbano Lorenzo-Seva 

Universitat Rovira i Virgili

ABSTRACT

A frequent criticism of exploratory factor analysis (EFA) is that it does not allow correlated residuals to be modelled, while they can be routinely specified in the confirmatory (CFA) model. In this article, we propose an EFA approach in which both the common factor solution and the residual matrix are unrestricted (i.e., the correlated residuals need not be specified a priori). The estimation procedures are two-stage and based on the unweighted least squares principle. Procedures for judging the solution appropriateness (including goodness of fit) are also proposed. The simulation studies and illustrative example suggest that the approach works quite well. Although the proposal is based on existing results, most of the developments can be considered to be new contributions, and are expected to be particularly useful in the earlier stages of item calibration. The whole procedure has been implemented in both R language and a well-known non-commercial EFA program.

KEYWORDS

Correlated residuals; exploratory factor analysis; item analysis; linear and nonlinear factor analysis; local dependence

A major criticism of unrestricted or exploratory factor analysis (EFA) with respect to its confirmatory counterpart (CFA) is that it does not allow correlated residuals to be modelled, while they can be routinely specified in CFA. This criticism, however, must be qualified: the “Classic” EFA assumption that the residual covariance matrix is diagonal (e.g., Mulaik, 2010) can be relaxed. And, in fact, ever since the 1960s more flexible forms of EFA that allow correlated residuals to be modelled have been proposed (Butler, 1968, McDonald, 1969, Mulaik, 2010, Yates, 1987). At present, correlated residuals within an EFA solution can also be modelled using ESEM (Asparouhov & Muthén, 2009, 2023, Heiserman & Maydeu-Olivares, 2017) or other more specific approaches based on the ESEM estimation procedure (*efast*; van Kesteren & Kievit, 2020, 2021). In these approaches, however, the residual structure is assumed to be known, so the correlated residuals that are to be estimated must be specified a priori (Asparouhov & Muthén, 2023, Van Kesteren & Kievit, 2021).

The topic of correlated residuals in EFA possibly originated from psychometric applications (Thurstone, 1947), but has permeated to other domains such as sociometrics and econometrics. (Blalock, 1971, Bollen, 1989, Costner & Schoenberg, 1973). The present proposal can also be used in these domains, but it is intended above all for item analysis applications in which correlated residuals have traditionally been referred to as “doublets” (Mulaik, 2010, Thurstone, 1947).

Overall, what we shall propose here is a simple and flexible approach for modelling an EFA solution with a

correlated-residual structure in which the nonzero correlated residuals do not need to be specified a priori. Therefore, both (a) the common-factor part, and (b) the residual correlation matrix of the solution that is fitted are unrestricted. This fully-unrestricted approach is expected to be useful when the practitioner (a) has reason to suppose that there are non-trivial residuals among the items (e.g., content redundancies); but (b) does not have enough information to specify which residuals are to be estimated. In other words, residuals are suspected, but how many and where they are located cannot be specified in advance. Substantively, this is the typical scenario found in the initial stages of test development, in which large pools of items with generally complex structures are analyzed so that those items that will constitute the final test version can be selected (Cattell, 1952, Ferrando et al., 2022, Floyd & Widaman, 1995; Marsh et al., 2014, Reise et al., 2000).

As a substantive basis for what we propose below, we shall now provide (a) a brief discussion about the “doublet controversy” (e.g., Asparouhov et al., 2015, MacCallum et al., 1992, Thurstone, 1947); and (b) some background on other proposals related to this one. As for the “doublet controversy”, if substantial correlated residuals are present in a data set and left unmodeled, two general problems are expected to appear. The first is bad model-data fit (Montoya & Edwards, 2021), which implies that additional non-content factors that account for the unmodeled redundancies will be needed to adequately reproduce the correlation matrix. The second is biased parameter estimates: loadings, residual variances and, in rotated solutions,

possibly also inter-factor correlations (Lorenzo-Seva & Ferrando, 2020, Mulaik, 2010, Van Kesteren & Kievit, 2021, Yates, 1987).

At the other extreme, however, blindly freeing pairs of residuals with the sole aim of achieving a good fit is expected to capitalize on chance (Browne, 2001). Goodness of fit will no doubt be improved in this particular sample, but possibly at the cost of modeling nonexistent residuals that are falsely detected because of sampling fluctuation. If this is the case, the solution obtained may be weak and unstable and poorly replicate under cross-validation. If biases are present, however, they will be far smaller than in this scenario (Mulaik, 2010, Reddy, 1992, Yates, 1987). So, the problem in this case is more of increased estimation error than of bias (e.g., DeMars, 2020).

As for background, the existing proposals based on fully unrestricted modeling (Butler, 1968, Mulaik, 2010, Yates, 1987) were developed not so much to interpret the parameters of a complete solution as to minimize the impact of the un-modeled doublets on the structural parameter assessment of the common part of the solution. So, essentially, the residual structure is viewed as a nuisance and a source of bias for the estimated loadings. In technical terms, these proposals were mainly descriptive, and not too concerned with either inferential issues (mainly standard errors and rigorous goodness of fit assessment) or the nature of the analyzed variables (e.g., approximately continuous or ordered categorical). In contrast, at the substantive level, the position we adopt here views the residual structure not as a nuisance, but as a source of relevant information. And, technically, our proposal is more complete, takes into account the nature of the variables and goes far beyond the descriptive aspects (see below). We would like to point out that the comparisons made here are only concerned with fully unrestricted solutions. If there is enough information to specify which doublets have to be estimated, the procedures mentioned above that require this information appear to work well and can be recommended (e.g., Van Kesteren & Kievit, 2021).

Although the present proposal is based on existing results, it does make three new contributions: (a) the development of an analytical procedure for determining which residuals are to be set to zero in the specification of the solution; (b) the full development of a limited-information extraction procedure for the specified solution that includes point estimates and standard errors for all the parameters; and (c) the adaptation of an empirical test statistic approach for this type of extended solution that allows the goodness of fit of the proposed solution to be rigorously assessed. At the instrumental level, the main contribution is that our proposal is implemented in both an R package and a widely known, user-friendly, and non-commercial EFA program.

1.1. General Model and Basic Results

The general starting model is an extension of a direct unrestricted multiple FA solution in which the initial, unrotated

pattern loading matrix is in canonical form (e.g., Harman, 1976). The extended model in the population is:

$$\Sigma = \Lambda\Lambda' + \Psi\Sigma_{uu}\Psi \quad (1)$$

where, Σ is the $n \times n$ inter-item correlation matrix, Λ is the $n \times r$ canonical unrotated pattern, Ψ is the $n \times n$ diagonal matrix containing the item residual standard deviations, and

Σ_{uu} is the $n \times n$ residual correlation matrix. In the classic EFA model, the residuals are all uncorrelated, so Σ_{uu} becomes an identity matrix, and model (1) reduces to

$$\Sigma = \Lambda\Lambda' + \Psi^2. \quad (2)$$

The structural model (1) can be applied to: (a) binary scores, (b) graded scores treated as ordered categorical variables, and (c) graded or more continuous scores treated as continuous variables. Depending on how the variables are treated, the type of elements in Σ (i.e., the inter-item correlations) in the general structure (1) will change: if the variables are treated as binary, the correlations will be tetrachoric; if treated as ordered-categorical, the correlations will be polychoric; and if treated as approximately continuous, the correlations will be product-moment (Pearson). Under normality assumptions and with reparameterization, the binary case of model (1) is a multidimensional two-parameter normal-ogive-model solution with correlated residuals, whereas the graded case is a multidimensional graded-response-model solution also with correlated residuals (see, e.g., McDonald, 1999, 2000).

The approach we propose below for fitting model (1) follows the same rationale as that for modelling the common-factor part of an exploratory FA solution based on the classic model (2). It is agreed that the “ideal” solution for the common part is a simple structure in which only a few loadings (in the appropriate places) are salient while the remaining are as close to zero as possible (e.g., McDonald, 2000). However, in an unrestricted solution, none of the loadings are fixed to zero (except, in some cases, those needed to identify the initial solution). If this rationale is extended to the correlated residual structure in (1), (a) the “ideal” solution for the residual correlation matrix is that in which only a few of its non-diagonal elements are truly bounded away from zero while the others are essentially zero (e.g., Pan et al., 2017), but, (b) only the minimal number of doublets needed for identification purposes are fixed to zero. If this solution is clear and approaches the ideal conditions above (both in terms of common-factor and residual structures) it can be used as a basis for specifying further, more restricted solutions either in new datasets or following a convincing cross-validation schema.

1.2. Parameter Estimation: The Two-Stage Approach

The general schema for fitting the extended EFA solution is a simple limited-information calibration-scoring procedure, which we consider to be quite suited to the conditions in which it is generally used. This paper discusses only the calibration stage of this schema (i.e., fitting the structural

solution (1)), and leaves the development of the scoring stage for a later proposal.

Using obvious notation, when fitted in the sample model (1) is written as:

$$\mathbf{R} = \mathbf{A}\mathbf{A}' + \mathbf{U}\mathbf{R}_{\mathbf{uu}}\mathbf{U} \quad (3)$$

The only additional assumption we require for estimating (1) from (3) is that the sample inter-item correlation matrix \mathbf{R} is positive definite (see Lorenzo-Seva & Ferrando, 2020).

1.2.1. First-Stage

The first stage aims to determine and estimate the elements of $\mathbf{\Sigma}_{\mathbf{uu}}$ that are the most substantial. To this end, the estimates of all the non-duplicated elements of $\mathbf{\Sigma}_{\mathbf{uu}}$ are first obtained by using the *residual omission* or *sectioning* procedure (Wright, 1968, Yates, 1987) developed by Ferrando et al. (2022). This procedure considers all pairs of non-duplicated elements to be possible doublets, and, for each pair, one element is omitted from the data set and the classical EFA model (2) is fitted to the remaining variables (the core set). This will provide initial estimates of the elements of $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ corresponding to the core variables. Next, the structural estimates for the variable not included in the core set are obtained by using extension analysis (e.g., Nagy et al., 2017). If we denote by $\mathbf{r}_{j,\text{core}}$ the column vector containing the correlations between variable j (omitted from the core set) and the variables in the core set, the pending estimates for the omitted variable are obtained as:

$$\mathbf{a}_j = (\mathbf{A}_{\text{core}}' \mathbf{A}_{\text{core}})^{-1} \mathbf{A}_{\text{core}}' \mathbf{r}_{j,\text{core}}; \mathbf{u}_j = \sqrt{1 - \mathbf{a}_j' \mathbf{a}_j} \quad (4)$$

If the procedure described above is applied sequentially to all pairs of non-repeated elements of $\mathbf{\Sigma}_{\mathbf{uu}}$, then, at the end of the process, initial estimates of $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ for the full set of variables in the analysis will have been obtained. Estimates of the residual correlations are now obtained by

$$\mathbf{R}_{\mathbf{uu}} = \mathbf{U}^{-1}(\mathbf{R} - \mathbf{A}\mathbf{A}')\mathbf{U}^{-1}. \quad (5)$$

The schema summarized so far is the same one that Ferrando et al. (2022) proposed for flagging potential doublets. In this previous proposal, once the estimates in (5) had been obtained, a re-sampling schema was developed to determine reference values for judging the significance of the elements of $\mathbf{\Sigma}_{\mathbf{uu}}$. So, Ferrando et al. (2022) propose only a detection procedure, which uses the point-estimates and accompanying reference values of the elements of $\mathbf{\Sigma}_{\mathbf{uu}}$ to provide a list of potential doublets. In the present proposal, however, we use result (5) as a starting point to develop a factor extraction procedure intended for fully unrestricted solutions of type (1), which includes: (a) parameter estimation (both point estimates and confidence intervals), and (b) goodness of model-data fit assessment. As far as we know, all the material below can be considered to be a new contribution.

Once the estimates of all the elements of $\mathbf{\Sigma}_{\mathbf{uu}}$ in (5) have been obtained, the next step is to determine which of them are (initially) "salient". To do so, the off-diagonal elements of $\mathbf{R}_{\mathbf{uu}}$ are arranged in descending order of absolute value and the first k elements are free. Now, if the value of k is

too large, there may not be enough degrees of freedom available to estimate and test the solution. For this reason, the maximum number of doublets that can be specified is limited to $k=g-r$, where g is Lederman's (1937) bound solved for the number of common factors (Hayashi & Marcoulides, 2006, Mulaik, 2010). The rationale for this limitation is discussed in Ferrando et al. (2022). By the end of the process, a trimmed estimate of $\mathbf{\Sigma}_{\mathbf{uu}}$ (denoted by $\mathbf{R}^{(t)}_{\mathbf{uu}}$) will have been obtained that has, at most, k nonzero elements. These nonzero estimates will be taken as fixed and known, and used in the second step described below.

It should be stressed that the selection process above does not imply that all the freely estimated k elements will necessarily be significantly different from zero in the population. Rather, they are just left free during the estimation process. As discussed above, the process applied to the k free elements in $\mathbf{R}_{\mathbf{uu}}$ is equivalent to that used in fitting the common structure of an exploratory solution: i.e., all loadings that do not need to be restricted for identification purposes are left free.

1.2.2. Second-Stage

The second stage uses only the $\mathbf{R}^{(t)}_{\mathbf{uu}}$ estimate obtained at the end of the previous stage. So, final estimates of $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ are now obtained based on $\mathbf{R}^{(t)}_{\mathbf{uu}}$. This stage consists essentially of fitting a conventional solution (2) to a reduced correlation matrix in which the first-order correlations in \mathbf{R} are partialled out using the residual structure determined in the first stage. More specifically, the final estimate \mathbf{U} of $\mathbf{\Psi}$ is first obtained as (see Mulaik, 2010)

$$\mathbf{U} = [\text{diag}(\mathbf{R}^{-1})]^{-1/2}. \quad (6)$$

Once this estimate has been obtained, the reduced matrix to be factored is:

$$\mathbf{R} - \mathbf{U}\mathbf{R}_{\mathbf{uu}}^{(t)}\mathbf{U} = \mathbf{A}\mathbf{A}' \quad (7)$$

and the factoring outcome on the right-hand side of (7) will provide the final estimates of $\mathbf{\Lambda}$. In principle, all the estimation procedures that are used to fit the classic model (2) can be adapted to fit (7). However, we shall only consider procedures based on the unweighted least squares (ULS) principle. ULS-EFA is easily implemented and computationally robust (Jöreskog, 2003, Knol & Berger, 1991 Lee et al., 2012, Forero et al., 2009, Zhang & Browne, 2006, Mislevy, 1986) which is particularly appropriate here given that model (1) can be far more parameterized (and so potentially unstable) than model (2).

Overall, the procedure proposed here can be viewed as an application of an ad-lib factorial process (Nunnally, 1978, p. 430) in which the estimates are obtained using various procedures (although they all follow the ULS principle). Thus, estimates in (5) are extension estimates (which can also be regarded as ULS estimates; see McDonald, 1978). The key point, however, is that the structural estimates in (7) are conditional upon the $\mathbf{R}^{(t)}_{\mathbf{uu}}$ estimates obtained in the first stage (which are taken as fixed and known without taking into account their uncertainty) and on the unicity estimates in (6).

So, while the procedure is expected to produce essentially unbiased estimates (Nunnally, 1978, Ten Berge, 1999), it cannot be claimed that they are efficient. What we can advance from now on, however, is that despite these limitations (or perhaps thanks to them), the procedure performs very well, particularly in the scenarios for which it is intended. Because the proposal is based on the correlated-residual estimates in (5), which belong to the MORGANA family of indices of this type (see Ferrando et al., 2022), we shall refer to the present method as Morgana Factor Analysis.

1.3. Assessing the Appropriateness of the Solutions

The appropriateness of a psychometric EFA solution must be assessed using a multifaceted approach that focuses on different groups of properties (e.g., Ferrando & Lorenzo-Seva, 2018). In the calibration stage, in particular, two main groups are important. The first is the degree of goodness of the model-data fit. The second is the strength and replicability of the structural solution obtained. The procedures for assessing this second group of properties do not depend on the estimation procedure and can be used directly with either of the two solutions (1) or (2). So, only goodness of fit (GOF) will be discussed here.

Because the estimation procedures we propose are far from efficient, no formal chi-square-based test of fit statistic can be easily derived in this case to judge the degree of model-data fit or to further obtain chi-squared-based goodness-of-fit indices. To address this limitation, we propose using the empirical test of fit statistic described in full in Lorenzo-Seva and Ferrando (2023) adapted to the present scenario. Only a summary of solution (1) is provided here.

The essential idea is to use the parent sample matrix \mathbf{R} in (3) as a basis for generating simulated pseudo-samples from a population in which the null hypothesis (1) holds (Bollen & Stine, 1992). Then, in each simulated pseudo-sample, the (correct) solution (7) is fitted to the corresponding matrix \mathbf{R}_i^* , and the ULS discrepancy function

$$c = (N - 1) \sum_{i=1}^{m-1} \sum_{j \neq i}^m e^2_{ij}, \quad (8)$$

is computed. The e^2_{ij} terms in (8) are the non-diagonal elements of the residual matrix:

$$\mathbf{E} = \mathbf{R} - \mathbf{U}\mathbf{R}_{\mathbf{U}\mathbf{U}}^{(t)}\mathbf{U} - \mathbf{A}\mathbf{A}'. \quad (9)$$

At the end of the sequence, the distribution of the pseudo-sample c values obtained when the null hypothesis holds is available. Because this distribution is not expected “per se” to be chi-squared, it is further transformed to be as close as possible. More specifically, a polynomial transformation is applied to the c values so that the transformed distribution has the first four moments of a chi-squared variable with degrees of freedom:

$$df = \frac{1}{2}(n - r)(n - r + 1) - n - nd. \quad (10)$$

where nd is the number of doublets that are freely estimated depending on the results obtained at the end of stage 1.

Conceptually, (10) is the standard number of degrees of freedom associated with an ULS unrestricted solution (e.g., Jöreskog, 1967, Lawley & Maxwell, 1973) with an additional loss of one degree of freedom for each freely estimated doublet.

Finally, once the null reference distribution is available, the observed c statistic is (a) obtained in the original sample using (8), (b) transformed using the same polynomial transformation, and (c) interpreted with relation to a chi-square distribution with degrees of freedom (10).

The empirical test described above is a test of exact fit, in which the null hypothesis assumes that solution (1) is correct (i.e., holds exactly in the population). However, as this hypothesis is patently false, rejecting it is just a matter of achieving enough power. To overcome this limitation, the strategy considered here is to use an approximate fit approach based on certain fit indices that are derived from the chi-squared statistic (MacCallum et al., 1996). Previous studies by Garrido et al. (2016), Yang and Xia (2015), as well as our experience suggest that the indices that work best with EFA solutions are the RMSEA and the CFI. So, here we shall propose to use these two indices derived from the test statistic in equations (8) to (10). Finally, and, as auxiliary measures of fit, we propose to use two indices that do not specifically depend on the estimation criteria and are based on the magnitude of the residuals: the RMSR and the GFI (e.g., Ferrando & Lorenzo-Seva, 2018, McDonald & Mok, 1995).

2. Simulation Studies

Three short simulation studies were planned to assess the performance of our approach. The first study assesses whether the loading estimates are correct and the ‘true’ correlated errors are properly detected. The second assesses whether the goodness-of-fit indices correctly indicate that the sample solution is compatible with the ‘true’ solution generated for the population. As the proposal is exploratory, this second study includes ‘true’ solutions with and without correlated errors. Finally, the third study assesses whether the goodness-of-fit indices correctly suggest that the sample solution is incompatible with the solution generated for the population when the sample specification does not match it.

2.1. First Study

A population loading matrix was defined using 10 variables that conformed a single-factor solution. The loading of each variable was uniformly taken from the range [.55-.70]. Two types of solution were defined. In the first type, all correlated errors were defined to be zero. In the second, three correlated errors were uniformly and randomly chosen from the range [.40-.50]. The sign of each correlated error was randomly chosen to be positive or negative. Based on the population solution, Monte Carlo simulation techniques were used to generate samples with $N = 1,000$.

For each sample, the Pearson correlation matrix was analyzed using (a) the classic model (2) fitted by ULS, and (b)

Morgana Factor Analysis in order to always retain a single factor. The bias between the population loading matrix and the sample loading matrices was estimated using the Root Mean Squared Discrepancy. We also recorded the number of times that the correlated error indices present in the population were defined by Morgana Factor Analysis as free correlated errors in each sample.

The study was replicated 500 times, and Morgana Factor Analysis converged to a proper solution 99.8% of the time. Table 1 shows the outcomes of the study. As can be observed, the classic-ULS solution produced some bias in the estimated loadings when correlated errors were present at the population level. However, Morgana Factor Analysis produced low biases in both situations. In addition, the 'true' correlated error indices were always correctly defined as free correlated errors in the analysis.

2.2. Second Simulation Study

The second simulation study focuses on the goodness-of-fit results when the model holds in the population. We aim to assess the performance of the fit indices both when there are correlated residuals in the population and when there are not.

A population loading matrix was defined using 10 variables that conformed a single factor. The loading value of each variable was uniformly taken from the range [.55-.70]. In addition, the following variables were manipulated:

1. Number of correlated residuals in the population solution. For each solution, a different number was chosen from zero to three.
2. When correlated errors were present, three levels of magnitude were defined: low [.20 - .30], medium

Table 1. Average bias in the loading matrix and results on the detection of correlated errors (standard deviations are provided in brackets).

	Classic-ULS-EFA		Morgana Factor Analysis	
	0	3	0	3
Number of correlated errors				
Bias	.0215 (.0053)	.0676 (.0111)	.0230 (.0056)	.0278 (.0062)
Percentage correct	–	–	–	100%

Table 2. Average of goodness-of-fit indices when correlated errors were present in the population (standard deviations are provided in brackets).

Method	Classic-ULS-EFA			Morgana Factor Analysis		
	RMSEA	CFI	TLI	RMSEA	CFI	TLI
OVERALL	.0940 (.0344)	.9007 (.0760)	.8724 (.0976)	.0047 (.0123)	.9976 (.0101)	.9968 (.0120)
N						
150	.0953 (.0342)	.8963 (.0714)	.8667 (.0918)	.0048 (.0133)	.9973 (.0065)	.9964 (.0098)
300	.0954 (.0327)	.8994 (.0711)	.8708 (.0909)	.0046 (.0122)	.9976 (.0062)	.9968 (.0094)
1000	.0912 (.0363)	.9064 (.0845)	.8796 (.1087)	.0048 (.0113)	.9978 (.0151)	.9973 (.0158)
Number of correlated errors						
1	.0694 (.0238)	.9466 (.0347)	.9313 (.0446)	.0026 (.0089)	.9983 (.0055)	.9979 (.0083)
2	.0963 (.0272)	.9008 (.0597)	.8725 (.0767)	.0046 (.0116)	.9978 (.0047)	.9971 (.0072)
3	.1162 (.0340)	.8547 (.0915)	.8133 (.1173)	.0069 (.0152)	.9967 (.0159)	.9956 (.0176)
Size of correlated errors						
Small	.0722 (.0236)	.9423 (.0366)	.9258 (.0470)	.0028 (.0086)	.9984 (.0033)	.9980 (.0050)
Medium	.0948 (.0285)	.9026 (.0598)	.8748 (.0769)	.0041 (.0113)	.9978 (.0051)	.9972 (.0078)
Large	.1149 (.0358)	.8572 (.0938)	.8165 (.1203)	.0073 (.0155)	.9965 (.0164)	.9953 (.0185)

Note: RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker and Lewis Index.

[.30 - .40], and large [.40 - .50]. The sign of each correlated error was randomly chosen to be positive or negative.

3. Sample size. For each population, Monte Carlo simulation techniques were used to produce samples of three sizes: 150, 300, and 1,000.

For each sample, the Pearson correlation matrix was analyzed using classic-ULS and Morgana Factor Analysis in order to always retain a single factor. The chi-square statistic was estimated as proposed in this paper. The goodness-of-fit indices based on the chi-square statistic were RMSEA, CFI, and TLI.

The study was replicated 500 times. The number of datasets analyzed when correlated errors were not present was 3 (sample sizes) \times 500 (replications) = 1,500. The number of datasets analyzed when correlated errors were present was 3 (sample sizes) \times 3 (number of correlated errors) \times 3 (magnitude of correlated errors) \times 500 (replications) = 13,500.

Morgana Factor analysis converged to a proper solution 13,471 times (99.8%) when correlated errors were present, and always when they were not.

Table 2 shows the outcomes when correlated errors were present. As can be observed, the goodness-of-fit indices for the classic ULS solutions revealed that the samples had a poor model fit, especially when the number of errors was three and the magnitude of correlated errors was large. In contrast, Morgana-derived goodness-of-fit indices consistently informed of a proper model fit in all situations.

Table 3 shows the outcomes when correlated errors were not present. Now, goodness-of-fit indices related to both the Classic ULS model and Morgana Factor Analysis informed of a proper model fit in all cases. As expected, the smaller the sample, the better the fit.

2.3. Third Simulation Study

Finally, the third study assessed whether the goodness-of-fit indices correctly suggest that the sample solution does not agree with the population solution in those cases that it does not.

A population loading matrix was defined using 15 variables that conformed an orthogonal three-factor solution, each factor defined by 5 variables. The salient loading value

Table 3. Average of goodness-of-fit indices when no correlated errors were present in the population (standard deviations are provided in brackets).

Method	Classic-ULS-EFA			Morgana Factor Analysis		
	RMSEA	CFI	TL	RMSEA	CFI	TL
OVERALL	.0614 (.0426)	.9433 (.0575)	.9271 (.0740)	.0026 (.0096)	.9982 (.0040)	.9977 (.0061)
N						
150	.0328 (.0279)	.9809 (.0214)	.9756 (.0277)	.0017 (.0073)	.9985 (.0031)	.9983 (.0047)
300	.0659 (.0312)	.9463 (.0376)	.9310 (.0484)	.0019 (.0077)	.9985 (.0031)	.9982 (.0048)
1,000	.0791 (.0360)	.9233 (.0524)	.9014 (.0673)	.0043 (.0127)	.9975 (.0053)	.9967 (.0081)

Note: RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker and Lewis Index.

Table 4. Average of goodness-of-fit indices when the model was wrongly specified.

Index	Classic ULS EFA	Morgana Factor Analysis
RMSEA	.1850 (.0648)	.0953 (.0372)
CFI	.4724 (.2673)	.8425 (.1066)
TLI	.4122 (.2810)	.7932 (.1472)

Note: RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker and Lewis Index.

of each variable was uniformly taken from the range [.55–.70], while the non-salient loading values were uniformly taken from the range [–.20–.20]. Three correlated errors were defined, which were uniformly and randomly chosen from the range [.40–.50]. The sign of each correlated error was randomly chosen to be positive or negative. Based on the population solution, Monte Carlo simulation techniques were used to produce samples with $N = 300$.

For each sample, the Pearson correlation matrix was analyzed using the Classic ULS solution and Morgana Factor Analysis so that a single factor was always retained (which was the wrong solution in this case). The GOF indices were the same as in the previous study.

The study was replicated 500 times, and Morgana Factor Analysis converged in a proper solution 99.7% of times. Table 4 shows the outcomes of the study. As can be observed, Classic-ULS goodness-of-fit indices reported a bad fit. The same indices derived from Morgana Factor Analysis also indicated that the fit was wrong. However, the latter suggested a better fit than those based on the classical solution. This result suggests that the misspecification of the solution was partly absorbed in the form of “inflated” correlated residuals. This problem, which is likely to appear in real applications, is discussed below in detail, but clearly shows that the proposal cannot be used uncritically.

3. Implementing the Proposal

All the procedures proposed here have been implemented in the 12.01 version of the program FACTOR (Ferrando & Lorenzo-Seva, 2018), a well-known non-commercial program for exploratory and semi-confirmatory FA. Furthermore, the Morgana-R program has been developed specifically for R users. The program is available at <https://www.psicologia.urv.cat/en/tools/morgana-r-code/>.

4. Illustrative Example

Although the procedure proposed here is expected to be particularly useful for large item sets and complex solutions, for

Table 5. Illustrative example. Item content.

1. Reading your horoscope is a superstition and a waste of time.
2. Horoscopes predict a person's future.
3. I like talking about subjects related to astrology.
4. The way I relate to other people depends on my zodiac sign.
5. It is essential that you know the zodiac sign of your boy(girl)friend.
6. Everybody should know their zodiac sign.
7. Astrology-based predictions are usually wrong.
8. Anything to do with astrology bores me.

the sake of clarity and for didactic and illustrative purposes, here we shall use a very simple and small example: a brief 8-item version of the Belief in Astrology Inventory (BAI; see Chico & Lorenzo-Seva, 2006 for details). We consider the example to be appropriate for highlighting the usefulness of the method. Indeed, the BAI items measure a unidimensional, conceptually narrow construct, which means that, a priori, the homogeneity of item content and, possibly, content redundancies are expected to be high (e.g., Reise & Waller, 2009). These features can easily be appraised by inspecting the item contents in Table 5. Our illustration re-analyses the data used in the original calibration of the inventory (Chico & Lorenzo-Seva, 2006). Participants were 743 undergraduates studying Psychology and Social Sciences at a Spanish university (84.1% females), aged between 18 and 60 years (mean: 21.7; standard deviation: 4.3).

Given that (a) the response format was ordered-categorical (5-point Likert); (b) some item distributions were asymmetrical (positively skewed); (c) the sample was large; and (d) the item set was small, we considered that the best choice for fitting the data was to use the nonlinear-FA model for ordered-categorical variables. So, as explained above, the analyses were based on the inter-item polychoric correlation matrix. As for data adequacy, the matrix was positive-definite, and showed a Kaiser-Meyer-Olkin (KMO) value of .824. Parallel Analysis suggested that a single factor should be retained (Timmerman & Lorenzo-Seva, 2011).

The data was first fitted using the classic EFA model in (2) with uncorrelated residuals by specifying a unidimensional solution. The estimation procedure was the same as the one used in the simulation studies above. The most important results concerning goodness of fit (GOF) are in the first row of Table 6 below

On the whole, the Classic EFA results above are compatible with an essentially unidimensional solution, but, at the same time, the fit cannot be considered to be good (the RMSR and RMSEA estimates are too high). However, inspection of the pattern (provided below) showed that essential unidimensionality was achieved as substantial loading estimates were obtained for all of the items, and the ECV

Table 6. Goodness of fit results for the unidimensional solutions with and without correlated residuals.

Model	χ^2 (df)	RMRS	RMSEA(90%CI)	CFI	GFI
Uncorrelated residuals	205.07(20)	.083	.112(.08; .12)	.994	.977
Correlated residuals	8.807(17)	.044	.001(.00; .005)	1.00	.985

Note: χ^2 = chi square; df = degrees of freedom; RMRS = Standardized Root Mean Squared Residual; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; GFI = Goodness of Fit Index.

Table 7. Pairs of items with freely estimated residuals (90%CI).

Pair	Estimated Residual Correlation Value
3. I like talking about subjects related to astrology. 8. Anything to do with astrology bores me.	-.524 (-.611; -.437)
5. It is essential that you know the zodiac sign of your boy(girl)friend. 6. Everybody should know their zodiac sing.	.465 (.414; .585)
5. It is essential that you know the zodiac sign of your boy(girl)friend. 8. Anything to do with astrology bores me.	-.024 (-.118; .141)

estimate was .784 (bootstrap 95% confidence interval .753 and .812). As for redundancies, at least two very large standardized residuals (estimates above 5) were observed. They involved the item pairs 3 and 8, and 5 and 6 (see below).

We turn now to the unidimensional version of solution (1) with correlated residuals fitted according to the procedure proposed in this paper. The GOF results are in the bottom row of Table 6. The doublets that were freely estimated by our procedure are in Table 7.

The fit results in Table 6 are quite clear. By all standards they are excellent and a great improvement on the classical solution. As for the pairs of residuals modeled by the procedure, the first two (3 and 8, and 5 and 6) seem to be substantial and can be justified. First, content inspection shows that redundancy is quite obvious in both cases: they are virtually opposite formulations of the same question. Second, they correspond to the largest standardized residuals flagged in the previous analysis. Third, the magnitudes of the correlations are substantial and the signs correspond to what could be expected given the item stems (see Table 5). In contrast, the third pair is far more questionable: the redundancy in content is not obvious, and the estimated correlation is quite low and does not reach significance (see the limits of the confidence interval). As discussed above, modelling an unnecessary doublet that is likely to be only noise is not expected to bias the structural parameters of the solution but adds unnecessary estimation error. So, this result clearly suggests further lines of action. First, a cross-validation study should be undertaken to see if the detection results generalize across different samples. If this is the case, which seems likely, this doublet should be left un-modeled in further studies based on more restricted solutions. More in detail, if a confirmatory, ESEM, or *efast* solution, is tried in a new sample, only the first two “solid” doublets need to be specified in advance.

Table 8 shows the loading pattern of the two competing solutions. The result is interesting and can be predicted in the simple scenario considered here (e.g., Costner & Schoenberg, 1973). When compared to the loading estimates obtained under the correlated-residual solution, the loadings of the items involved in the doublets (3,5,6, and 8) are inflated (upwardly biased) while the estimates for the remaining items are deflated (downwardly biased). So,

Table 8. Loading pattern corresponding to the two solutions.

Item	Uncorrelated residuals	Correlated residuals
1	-.638	-.669
2	.435	.456
3	.605*	.539
4	.694	.730
5	.715*	.633
6	.697*	.647
7	-.525	-.554
8	-.658*	-.559

Note: * loadings of the items involved in doublets.

differential bias effects due to the presence of unmodeled doublets can be clearly observed in Table 8.

5. Discussion and Conclusions

The classical EFA model assumes that the variables under study are no longer correlated once the effect of the prescribed common factors are partialled out. In item analysis, however, correlated residuals are very common in most applications, particularly at earlier stages of the analysis. When this occurs, classical EFA is simply a “wrong” model, and forcing the data to fit it is only expected to lead to distorted results, particularly in terms of biased parameter estimates and incorrect goodness of fit assessment. The present simulation results clearly illustrate this point.

In this article we have proposed a simple approach for fitting a fully unrestricted EFA solution that incorporates correlated residuals. So, the solutions to be fitted do not need to specify the residuals that should be freed. This approach contrasts with existing proposals in which the common-factor part is unrestricted but the residual structure is specified a priori (which implies that more information is available for the residual structure than for the common structure). This scenario is justified in some settings (e.g., Van Kesteren & Kievit, 2021) but, in our opinion, not in the first stages of item analysis. Rather, the typical scenario in this case is that: (a) information about the common structure is available (although not to the point to which a restricted solution can be specified) and (b) redundancies in the form of doublets are strongly suspected but how many there are and where they are located cannot

be specified in advance. In this type of setting, our proposal addresses the needs of applied research. Furthermore, it is flexible and as simple as possible, properties that are particularly valued in situations where there are a large number of variables and samples that are not too large.

The simulation results and the illustrative example suggest that the proposal works quite well in spite of this simplicity. In the conditions considered, it always produced virtually unbiased estimates and correct goodness of fit results. The only negative result is a certain tendency to overfit (see below), which is only to be expected when the flexibility of any factorial solution is increased.

In spite of the promising results so far, it is indeed acknowledged that our proposal has its share of limitations and issues that require further research. Thus, more extensive simulations and empirical studies should be undertaken to ascertain its appropriateness in certain scenarios. Also, the goodness-of-fit proposals associated with the procedure are only tentative and should be further assessed. Finally, only the calibration (structural) stage has been considered in the proposal. Factor-scoring in the presence of correlated residuals is the next topic that should be studied.

Any methodological proposal can be misused, and this is also the case here. So, two main initial caveats are in order. First, Morgana EFA is intended only as a first step before more refined and restricted solutions are fitted based on cross-validation. Second, it is not intended to be used blindly and uncritically. The following two likely scenarios show the dangers of doing so. In the first, the procedure frees doublets that are not substantial (or even significant). This is capitalization on chance, and may lead to overfitting and potential instability. In the second scenario, the researcher specifies an insufficient number of common factors, which means that some “content” common variance is still left unmodeled. If this occurs, this variance is expected to be absorbed in the residual correlation matrix, so that the elements of this matrix would now reflect a mixture of shared specificities and unmodeled common variance. Note that this is the reverse of the “propagation” problem that occurs when ‘true’ doublets are left unmodeled, in which the residual correlations are “absorbed” by the common structural parameter estimates. In the first scenario, the researcher would be well advised to check both the meaning and the significance of the doublet estimates and fit further simplified solutions if appropriate. In the second, they should try solutions with different number of specified common factors, check the outcomes in terms of estimated doublets, and examine their meaning and content (i.e., plausible doublets vs. a true conglomerate of variables that probably defines a content factor). It would also be useful to assess the number of shared consistencies in residual matrix (9), for example by fitting an “extra” common factor and/or computing an index such as the KMO. If there are a considerable number of shared consistencies, a re-specified solution with more common factors should at least be tried. To sum up, when researchers use Morgana EFA there is nothing that exempts the researcher from thinking and deciding critically.

In spite of the limitations and the dangers discussed above, we believe that, if used correctly, what we propose here will be of great help to the applied researcher. It is a simple and efficient tool that has wide applicability and a clear rationale. Furthermore, everything we have proposed here is implemented in a non-commercial widely known program, as well as in an R package. So, it can be used from now on in any application that requires it.



Acknowledgements

This project has been made possible by the support of the Ministerio de Ciencia e Innovación, the Agencia Estatal de Investigación (AEI) and the European Regional Development Fund (ERDF) (PID2020-112894GB-I00).

Ethics Statement

The research has been approved by *The Ethics Committee Concerning Research into People, Society and the Environment* of the *Universitat Rovira i Virgili*. Report number CEIPSA-2021-PR-0028.

ORCID

Pere J. Ferrando  <http://orcid.org/0000-0002-3133-5466>
 Ana Hernández-Dorado  <http://orcid.org/0000-0001-9502-9735>
 Urbano Lorenzo-Seva  <http://orcid.org/0000-0001-5369-3099>

References

- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16, 397–438. <https://doi.org/10.1080/10705510903008204>
- Asparouhov, T., & Muthén, B. (2023). Residual structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 30, 1–31. <https://doi.org/10.1080/10705511.2022.2074422>
- Asparouhov, T., Muthén, B. y., & Morin, A. J. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeier. *Journal of Management*, 41, 1561–1577. <https://doi.org/10.1177/0149206315591075>
- Blalock, J. (Ed.) (1971). *Causal models in the social sciences*. Macmillan press. <https://doi.org/10.4324/9781315081663>
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17, 303–316. <https://doi.org/10.1177/0049124189017003004>
- Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, 21, 205–229. <https://doi.org/10.1177/0049124192021002004>
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111–150. https://doi.org/10.1207/S15327906MBR3601_05
- Butler, J. M. (1968). Descriptive factor analysis. *Multivariate Behavioral Research*, 3, 355–370. https://doi.org/10.1207/s15327906mbr0303_5
- Cattell, R. B. (1952). *Factor analysis: An introduction and manual for the psychologist and social scientist*. Harper.
- Chico, E., & Lorenzo-Seva, U. (2006). Belief in astrology inventory: Development and validation. *Psychological Reports*, 99, 851–863. <https://doi.org/10.2466/PRO.99.3.851-863>
- Costner, H., & Schoenberg, R. (1973). Diagnosing indicator ills in multiple indicator models. In A. S. Goldberger, and O. D. Duncan (Eds.), *Structural equation models in the social sciences*. Seminar (Seminar press). (pp. 167–499).
- DeMars, C. E. (2020). Comparing Causes of Dependency: Shared Latent Trait or Dependence on Observed Response. *Journal of Applied Measurement*, 21, 400–419.

- Ferrando, P. J., Hernandez-Dorado, A., & Lorenzo-Seva, U. (2022). Detecting correlated residuals in exploratory factor analysis: New proposals and a comparison of procedures. *Structural Equation Modeling: A Multidisciplinary Journal*, 29, 630–638. <https://doi.org/10.1080/10705511.2021.2004543>
- Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, 78, 762–780. <https://doi.org/10.1177/0013164417719308>
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286–299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 16, 625–641. <https://doi.org/10.1080/10705510903203573>
- Garrido, L. E., Abad, F. J., & Ponsoda, V. (2016). Are fit indices really fit to estimate the number of factors with categorical variables? Some cautionary findings via Monte Carlo simulation. *Psychological Methods*, 21, 93–111. <https://doi.org/10.1037/met0000064>
- Hayashi, K., & Marcoulides, G. A. (2006). Teacher's corner: Examining identification issues in factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 13, 631–645. https://doi.org/10.1207/s15328007sem1304_7
- Heiserman, N., & Maydeu-Olivares, A. (2017). *Best practices for exploratory factor analysis: Target rotations and correlated errors*. Columbia.
- Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, 32, 443–482. <https://doi.org/10.1007/BF02289658>
- Jöreskog, K. G. (2003). Factor analysis by MINRES. *To the Memory of Harry Harman and Henry Kaiser*. https://www.ssicentral.com/wp-content/uploads/2020/07/lis_minres.pdf
- Knol, D. L., & Berger, M. P. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research*, 26, 457–477. https://doi.org/10.1207/s15327906mbr2603_5
- Lawley, D. N., & Maxwell, A. E. (1973). Regression ana factor analysis. *Biometrika*, 60, 331–338. <https://doi.org/10.1093/biomet/60.2.331>
- Ledermann, W. (1937). On the rank of the reduced correlational matrix in multiple-factor analysis. *Psychometrika*, 2, 85–93. <https://doi.org/10.1007/BF02288062>
- Lee, C. T., Zhang, G., & Edwards, M. C. (2012). Ordinary least squares estimation of parameters in exploratory factor analysis with ordinal data. *Multivariate Behavioral Research*, 47, 314–339. <https://doi.org/10.1080/00273171.2012.658340>
- Lorenzo-Seva, U., & Ferrando, P. J. (2020). Unrestricted factor analysis of multidimensional test items based on an objectively refined target matrix. *Behavior Research Methods*, 52, 116–130. <https://doi.org/10.3758/s13428-019-01209-1>
- Lorenzo-Seva, U., & Ferrando, P. J. (2023). A simulation-based scaled test statistic for assessing model-data fit in least-squares unrestricted factor-analysis solutions. *Methodology*, 19, 96–115. <https://doi.org/10.5964/meth.9839>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1, 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490–504. <https://doi.org/10.1037/0033-2909.111.3.490>
- Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, 10, 85–110. <https://doi.org/10.1146/annurev-clinpsy-032813-153700>
- McDonald, R. P. (1969). A generalized common factor analysis based on residual covariance matrices of prescribed structure. *British Journal of Mathematical and Statistical Psychology*, 22, 149–163. <https://doi.org/10.1111/j.2044-8317.1969.tb00427.x>
- McDonald, R. P. (1978). McDonald, R. P. (1978) Some checking procedures for extension analysis. *Multivariate Behavioral Research*, 13, 319–325. https://doi.org/10.1207/s15327906mbr1303_4
- McDonald, R. P. (1999). *Test theory: A unified treatment*. psychology press.
- McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, 24, 99–114. <https://doi.org/10.1177/01466210022031552>
- McDonald, R. P., & Mok, M. M. C. (1995). Goodness of fit in item response models. *Multivariate Behavioral Research*, 30, 23–40. https://doi.org/10.1207/s15327906mbr3001_2
- Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics*, 11, 3–31. <https://doi.org/10.3102/10769986011001003>
- Montoya, A. K., & Edwards, M. C. (2021). The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. *Educational and Psychological Measurement*, 81, 413–440. <https://doi.org/10.1177/0013164420942899>
- Mulaik, S. A. (2010). *Foundations of factor analysis*. (2nd ed.). CRC Press. <https://doi.org/10.1201/b15851>
- Muthén, B. O. (1993). Goodness of fit test with categorical and other nonnormal variables. In K. A. Bollen and J. S. Long (Eds.), *Testing structural equation models*. (pp. 205–234). SAGE.
- Nagy, G., Brunner, M., Lüdtke, O., & Greiff, S. (2017). Extension procedures for confirmatory factor analysis. *The Journal of Experimental Education*, 85, 574–596. <https://doi.org/10.1080/00220973.2016.1260524>
- Nunnally, J. C. (1978). *Psychometric theory*. 2nd Edition, McGraw-Hill.
- Pan, J., Ip, E. H., & Dubé, L. (2017). An alternative to post hoc model modification in confirmatory factor analysis: The Bayesian lasso. *Psychological Methods*, 22, 687–704. <https://doi.org/10.1037/met0000112>
- Reddy, S. K. (1992). Effects of ignoring correlated measurement error in structural equation models. *Educational and Psychological Measurement*, 52, 549–570. <https://doi.org/10.1177/0013164492052003005>
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, 5, 27–48. <https://doi.org/10.1146/annurev.clinpsy.032408.153553>
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12, 287–297. <https://doi.org/10.1037/1040-3590.12.3.287>
- Ten Berge, J. M. (1999). A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research*, 34, 89–102. https://doi.org/10.1207/s15327906mbr3401_4
- Thurstone, L. L. (1947). *Multiple-factor analysis; a development and expansion of the vectors of mind*. University of Chicago Press.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16, 209–220. <https://doi.org/10.1037/a0023353>
- van Kesteren, E.-J., & Kievit, R. A. (2020). efastGitHub. <https://doi.org/10.5281/zenodo.3779927>
- Van Kesteren, E.-J., & Kievit, R. A. (2021). Exploratory factor analysis with structured residuals for brain network data. *Network Neuroscience (Cambridge, Mass.)*, 5, 1–27. https://doi.org/10.1162/netn_a_00162
- Yang, Y., & Xia, Y. (2015). On the number of factors to retain in exploratory factor analysis for ordered categorical data. *Behavior Research Methods*, 47, 756–772. <https://doi.org/10.3758/s13428-014-0499-2>
- Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. State University of New York Press.
- Zhang, G., & Browne, M. W. (2006). Bootstrap fit testing, confidence intervals, and standard error estimation in the factor analysis of polychoric correlation matrices. *Behaviormetrika*, 33, 61–74. <https://doi.org/10.2333/bhmk.33.61>