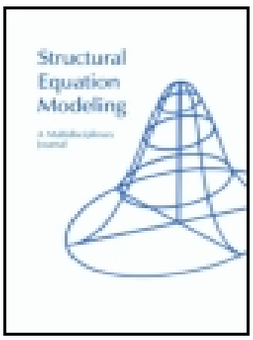

Routledge
Taylor & Francis Group

Structural Equation Modeling: A Multidisciplinary Journal

ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: https://www.tandfonline.com/loi/hsem20

Not Positive Definite Correlation Matrices in Exploratory Item Factor Analysis: Causes, Consequences and a Proposed Solution

Urbano Lorenzo-Seva & Pere J. Ferrando

To cite this article: Urbano Lorenzo-Seva & Pere J. Ferrando (2020): Not Positive Definite Correlation Matrices in Exploratory Item Factor Analysis: Causes, Consequences and a Proposed Solution, Structural Equation Modeling: A Multidisciplinary Journal, DOI: 10.1080/10705511.2020.1735393

To link to this article: https://doi.org/10.1080/10705511.2020.1735393

Published online: 13 Mar 2020.

STRUCTURAL EQUATION MODELING: A MULTIDISCIPLINARY JOURNAL
https://doi.org/10.1080/10705511.2020.1735393

Routledge
Taylor & Francis Group

TEACHER'S CORNER

OPEN ACCESS Check for updates

Not Positive Definite Correlation Matrices in Exploratory Item Factor Analysis: Causes, Consequences and a Proposed Solution

Urbano Lorenzo-Seva and Pere J. Ferrando

Rovira i Virgili University (SPAIN)

ABSTRACT

Least-squares exploratory factor analysis based on tetrachoric/polychoric correlations is a robust, defensible and widely used approach for performing item analysis, especially in the first stages of scale development. A relatively common problem in this scenario, however, is that the inter-item correlation matrix fails to be positive definite. This paper, which is largely intended for practitioners, aims to provide a didactic discussion about the causes, consequences and remedies of this problem. The discussion is more applied than statistical and based on the factor analysis model, and the problem is linked to that of improper solutions. Solutions for preventing the problem from occurring, and the smoothing corrections available at present are described and discussed. A new smoothing algorithm is also proposed.

KEYWORDS

Not positive definite; polychoric correlation; smoothing method; Heywood cases

The exploratory or unrestricted factor analysis (EFA) model continues to play an important role in the development, validation and usage of most psychometric measures, particularly in the non-cognitive or typical-response domains (e.g. Reise, Waller, & Comrey, 2000). In the first stages of the development of a measure, large item pools are usually analyzed to determine the most appropriate dimensionality and structure of the item scores. In this scenario, EFA is particularly appropriate for assessing these basic issues. Information gained at further stages, when 'cleaned' item pools are assessed, might allow more restricted solutions to be tested with a confirmatory model. However, the inherent complexity of many typical-response items (e.g. Cattell, 1986) makes the more flexible EFA model a potentially useful option even in this case (Ferrando & Lorenzo-Seva, 2000).

The overwhelming majority of items to which the EFA model is applied use ordered-categorical response formats, which typically range between 2 and 7 categories. Scores produced by these items can be treated (a) as approximately continuous measures by using the standard linear factor analysis (FA) model; or (b) as ordinal measures by using Categorical-Variable (CV) FA, which is generally based on an underlying-variables approach (UVA, e.g. Muthén, 1984, 1993). Essentially, the UVA approach assumes that continuous latent variables of response strength underlie the observed categorical variables, so that these categorical variables result from the categorization of the underlying variables at given thresholds.

Considerable debate exists regarding the appropriateness and usefulness of one approach or another, and we shall only further discuss this point in terms of its relevance to this article. At present, it is accepted that CV-UVA EFA based on categorical item scores is an approach that in most cases mitigates the problems of spurious evidence of multidimensionality and differentially attenuated loading estimates which arise because of the nonlinearity in the item-factor regressions (McDonald & Ahlawat, 1974). The number of practitioners who use this approach in item analysis is also clearly increasing: a research in google academic of the terms "polychoric" and "exploratory factor analysis" produced an outcome of 3,810 items until the year 2009. However, the same search in the period 2010–2019 produced an outcome of 12,500 items. However, this increase in use has also raised awareness among practitioners that CV-UVA EFA is not devoid of problems either.

In its most basic form, the CV-UVA entails fitting the common EFA model to the tetrachoric (binary) or polychoric (polytomous) inter-item correlation matrix. Because the tetrachoric correlation is the particular case of the polychoric correlation when there are only two categories, from now on we shall use the term polychoric to refer to both. The basic approach we are describing is usually known as limited-information estimation of the item EFA model. Its main alternative are the full-information methods, that use the raw data directly in parameter estimation, and so, avoid the intermediary step of obtaining the inter-item polychoric matrix (e.g. Wirth & Edwards, 2007). Full-information methods are theoretically superior than limited-information methods, and a great effort on improving its functioning is being made from the item response theory framework. For the moment, however, and in the context of item EFA, the limited-information approach works as well in practice, is computationally superior, especially as the number of factors increase, and procedures for assessing model-data fit are much more developed in this approach (e.g. Barendse, Oort, & Timmerman, 2015). Finally, an additional advantage of limited-information item EFA is that the model can be extended to multiple-group and longitudinal analyses, as

CONTACT Urbano Lorenzo-Seva urbano.lorenzo@urv.cat Facultat de Ciències de l'Educació i Psicologia, Universitat Rovira i Virgili, Carretera Valls s/n, Tarragona 43007, Spain.

well as full structural models in general. These types of extensions are generally known as ESEM (Asparouhov & Muthén, 2009).

It must be noted that the polychoric correlation is not a statistic obtained directly from the data, but an estimator of a latent correlation between assumed continuous response variables that is estimated iteratively and which is quite complex. This estimator may not converge, achieve implausible results, or may simply be very imprecise, with typical errors much larger than those of a Pearson correlation. Among other factors, these problems depend on the sample size and the number of response categories (Ferrando & Lorenzo-Seva, 2014). Underlying non-normality induces bias in polychoric estimates and their standard errors: the bias also affects to model estimates and goodness-of-fit tests (Foldnes & Grønneberg, 2019). In addition, model misspecification can lead to sample correlation estimates notably different from the population polychoric correlations and as a result, the reliability estimate larger than 1 (Kim, Lu, & Cohen, 2020).

Now, one of the potential problems found when a polychoric matrix is factor-analyzed is that it fails to be positive definite. As developers of free software for exploratory item factor analysis, we can attest that, in this domain, the problem is relatively frequent, and is also a cause of considerable uneasiness and concern among practitioners (Is the data wrong?; Is the model wrong?; Have I done something wrong?). A review of the 307 queries received over more than 10 years is presented in Table 1.

While the most frequent questions are on: (a) the comparison between the outcomes of methods (like Parallel Analysis versus Hull, or polychoric versus Pearson correlations), (b) the interpretation of the EFA solution obtained in a particular dataset, (c) which indices are actually implemented (like the indices related to the reliability of factor scores), or (d) how to format a datafile so that it can be analyzed with our software, the questions on not positive definite correlation matrices are the ones which most concern researchers because they do not understand the surprising results they have obtained. It must be noted that most researchers probably do not even realize that they have run into a correlation matrix that fails to be positive definite when the negative eigenvalues are actually close to zero, and their impact on the factor solution is minor. However, they seek our advice when they face much more difficult situations. For example, one researcher obtained a tetrachoric correlation matrix between 36 variables that have six negative eigenvalues, the largest of which had a value of −0.305. In such extreme situations, the factor solution is so distorted that researchers cannot determine what has happened, and this is when they ask for our advice.

Table 1. Frequency of queries addressed to the authors on different aspects of exploratory factor analysis.

Source of query	Frequency
Parallel analysis, MAP and HULL	15%
Interpretation of model	13%
Data file format	12%
Technical questions, bugs report	12%
Polychoric vs Pearson correlations	9%
Reliability	8%
Not positive definite correlation matrix	7%
Goodness-of-fit indices	6%
Factor extraction methods	3%
Suggestions for further improvements	3%
Handling of missing values	3%
Rotation methods	3%
Factor scores	2%
Unidimensionality	2%
Others	2%
Sample size	1%
Bifactor	1%

The present article has a dual purpose: didactic and methodological. First, we aim to provide a clear discussion of the problem, including causes, consequences and remedies, which can serve as a useful guide for practitioners that use EFA for psychometric purposes. Second, we propose a new approach for smoothing indefinite polychoric inter-item matrices, with minimal impact on the smoothed correlation matrix. With regards to the first contribution, we agree with Wothke (1993) that the clear mathematical definition of positive definitiveness in itself is generally of little help to practitioners. So we have provided a detailed and specific treatment in terms of the EFA model, linking the failure to meet these conditions to the problems of improper or inadmissible solutions, because these should be, in our opinion, the main causes of concern when the problem occurs. As for the second, methodological contribution, our proposal is a refinement of a previous methodology developed by Bentler and Yuan (2011). As the short simulation below suggests, our proposal seems to improve the behavior of the original proposal on which is based.

Basic theoretical results

An inter-item correlation matrix is positive definite (PD) if all of its eigenvalues are positive. It is positive semidefinite (PSD) if some of its eigenvalues are zero and the rest are positive. Finally, it is indefinite if it has both positive and negative eigenvalues (e.g. Wothke, 1993). This last situation is also known as not positive definite (NPD).

The eigenvalues of the inter-item correlation matrix can be interpreted as the amounts of total variance explained by the corresponding principal components of the items. On the other hand, the eigenvalues of the reduced inter-item correlation matrix with proper communalities in the main diagonal can be interpreted as the amounts of common variance explained by the corresponding common factors (see for example, ten Berge, 1998). Under standard conditions (discussed below), the number of principal components which are needed to account for all the total variance is the same as the number of items. In the EFA model, however, the number of common factors needed to account for all the common variance is assumed to be less than the number of items. Finally, because variances (sums of squares) are positive quantities, negative amounts of total or common variances (i.e. negative eigenvalues) are inadmissible results that make no sense.

From the summary provided above, it should be clear that in the EFA model, and under standard conditions, the inter-item correlation matrix in the population is PD, and the reduced correlation matrix in the population is PSD (because

the number of common factors needed to explain all the common variance is less than the number of items, some of the eigenvalues of the reduced matrix must be zero). So, the sample correlation matrix (denoted as **R** from now on) is an estimate of a PD population matrix, and the sample reduced matrix with communality estimates on the main diagonal (denoted as **R*** from now on) is an estimate of a PSD population matrix. In some circumstances, however, both sample matrices can fail to meet these requirements.

Sources of not positive definiteness

We shall start by considering the case of product-moment (Pearson) inter-item **R**s. In principle, this type of matrix must be PD if some standard conditions are met (e.g. Gorsuch, 1983; Wothke, 1993). Of these conditions, four are particularly important for the present developments. First, the number of observations from which **R** is computed is larger than the number of items. Second, all the correlations are based on the same number of cases. Third, there are no linear dependencies among the item scores (i.e. certain inter-item correlations are unit or near unit, or certain item scores are linear composites of the remaining scores). Fourth, there are no items with zero variance. The second condition is violated when correlations are obtained under pairwise deletion because, if they are, each correlation is obtained from a different sub-set of cases (e.g. Arbuckle, 1996; Wothke, 1993). The third condition will not be met when the same item is repeatedly entered into the data set, or if the dataset contains highly redundant items that correlate near one among them. Finally, the fourth condition will not be met if extreme items or, more generally, items for which all respondents provide the same answer are present. Note that these conditions are basic design requirements that do not depend on the type of correlation matrix (i.e., a Pearson correlation matrix or a polychoric correlation) that will be factor analyzed. However, if these conditions are not met when polychoric matrices are factor analyzed, the problems this article deals with will probably be aggravated. In our experience, most of the inter-item polychoric matrices reported to be NPD also suffer from some of the design problems described above.

So, in principle, two basic recommendations need to be made. The first one is to use listwise computation and, if this is not possible, schemas other than pairwise deletion should be used to deal with missing data (see Arbuckle, 1996; Lorenzo-Seva & Van Ginkel, 2016; Wothke, 1993). The second one is to 'clean' the data set and remove the offending items described above, which, in addition, do not provide any relevant information. This initial screening can be tedious in large item sets but is well worth the trouble.

We turn now to inter-item polychoric matrices. Although joint estimation of the full matrix is possible (e.g. Lee & Poon, 1987), for practical reasons the elements of these matrices are generally estimated on a bivariate or pairwise basis (joint estimation becomes computationally intractable with a large number of items), and this is certainly the case with EFAs based on many items. Now, pairwise estimation does not ensure positive definiteness of **R**. So, it may or may not be PD even if the standard conditions above are met.

As explained above, in the EFA model the polychoric **R** is an estimate of a PD population matrix. Furthermore, each polychoric correlation obtained in the sample is a consistent estimate of its population counterpart. So, if the model is correct, as the sample size increases, each element of **R** will increasingly approach its corresponding element in a PD population matrix. It then follows that, if the standard preliminary conditions above are met, the problem of **R** being NPD is a problem of sampling fluctuation, so the conditions that make this problem more or less likely to appear can be determined. These conditions are: (a) sample size, (b) number of items, (c) number of response categories, (d) item extremeness, and (e) magnitude of the inter-item correlations.

Conditions (a) and (b) above are clear, and match the empirical evidence: most **R** matrices reported to be NPD were based on large item sets administered in medium-to-small samples. Please note that small sample size means large sampling error, so the sample polychoric correlation estimate can be very different from the population element it estimates. And with a large number of items there is more room for some elements of **R** to considerably depart from their population counterparts.

Conditions (c) to (e) above are, perhaps, less immediate. As for (c), if the thresholds corresponding to the categories are taken as fixed, and other things remain constant, the sampling error of a polychoric correlation decreases with the number of categories. So, the tetrachoric estimate based only on two categories is the one that has maximal sampling error (e.g. Guilford & Lyons, 1942, Olsson, 1979). And, as for (d), the sampling error of a polychoric estimate varies widely as a function of item locations or thresholds, so it increases as the items become more extreme (e.g. Guilford & Lyons, 1942; Mislevy, 1986; Olsson, 1979). It then follows that large sampling errors are expected in item sets with a wide spread of extremeness or locations. Finally, (e) (the role of the magnitude of the correlations, which determines the internal consistency of the item set) impacts the NPD outcome in two opposite ways: all other things being equal, the sampling error of a polychoric correlation decreases as the 'true' correlation value increases (Olsson, 1979; we assume for the sake of simplicity that the correlation is positive). However, if the 'true' correlation value is near its upper limit it is more likely to lead to communality estimates above the unit upper limit. So, in summary, the failure of **R** to be PD is mostly expected in the case of a large set of binary items that vary widely in extremeness, some of which are highly correlated (in the form of redundant content, doublets or triplets), and are administered in a small sample.

Consequences of not positive definiteness in EFA

In many cases, EFA can still be directly computed when **R** fails to be PD. Even in this case, however, several problems of estimation, testing, and interpretation of the results might appear. In general, the feasibility of EFA as well as the potential problems depend on (a) the extraction criterion on which the EFA is based, (b) the particular estimation technique used to meet this criterion, and (c) the relevance of the not positive definiteness.

We shall start with the two preliminary checks that are commonly used in EFA to assess whether the dataset is suitable for fitting the model to it: Bartlett's (1950) test and the Kaiser, Meyer and Olkin measure of sampling adequacy (KMO; Kaiser, 1970; Kaiser & Rice, 1974). If one or more of the eigenvalues of **R** are zero, the matrix will be singular and cannot be inverted, which implies that neither of the two indices can be obtained. If some of the eigenvalues of **R** are negative, either no results can be obtained, or these results are likely to be implausible (e.g. results outside the 0–1 theoretical bounds in the KMO case). These are the first "odd" results practitioners generally obtain when they try to fit the EFA model and **R** is NPD.

We turn now to fitting the EFA model in the strict sense (i.e., to compute the estimates of the model parameters from the sample data). In general terms, some procedures such as generalized least squares (GLS) or minimum–rank FA (ten Berge & Kiers, 1991) are simply not feasible when **R** is NPD. Other procedures such as Maximum Likelihood (ML) EFA are feasible in some cases. However, even when the analysis can be performed, problems of convergence, unstable estimates and unstable goodness-of-fit measures can be expected when **R** fails to be PD and ML-FA is used (Yuan, Wu, & Bentler, 2011).

The criterion in which positive definiteness of **R** is not so important (see e.g. Wothke, 1993) is unweighted (or ordinary) least squares (ULS), and from now on we shall solely consider this approach. First, it seems to be of more interest to study the consequences of the not positive definiteness of **R** precisely in the case in which the analysis is most likely to be able to converge in a solution. Second, these consequences are particularly clear in this case. So, ULS-EFA is submitted as the best choice for didactic purposes. Third, ULS-EFA is particularly appropriate in the scenarios considered here: It is easily implemented, computationally robust, and works well in large item sets and samples that are not too large (Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009; Lee, Zhang, & Edwards, 2012; Mislevy, 1986; Zhang & Browne, 2006). In fact, earlier studies (Knol & Berger, 1991) suggest that the simple ULS-EFA based on polychoric **R** provides better estimates and measures of fit than methods that are either more theoretically correct or use more information from the data or both.

Several EFA procedures and numerical solutions are based on the ULS criterion and have been given different names in the literature. In general, however, they can all be categorized into two main approaches (Harman & Jones, 1966; McDonald, 1985, p. 94). In the first approach, a reduced $\mathbf{R}^{\star}$ is obtained by using communality estimates, which may be fixed from the start or iterated until minimum function values are obtained, and the ULS solution is obtained via principal axis factoring of $\mathbf{R}^{\star}$. In the second approach, pattern loading estimates that directly minimize the ULS function are obtained by using approximation procedures, and the communalities are then obtained on the basis of the loading estimates of each approximation. Harman's MINRES is possibly the best known version of this second approach (Harman & Jones, 1966). In principle, in none of the two ULS approaches above is **R** required to be PD. However, when it fails to be, some consequences, which are the same for both approaches, are expected. Interested readers can obtain a formal discussion of this by e-mail, or download it from the website http://psico.fcep.urv.cat/utilitats/factor/Documentation.html. Here we shall provide a conceptual discussion.

The general consequences of **R** being not PD are improper solutions in which one or more estimates have inadmissible values (i.e. values outside their theoretical bounds; see Bentler & Yuan, 2011). These inadmissible estimates are (a) uniqueness with zero or negative values, or (b) equivalently unit or greater than one communalities. In the FA literature these inadmissible results are known as Heywood (1931) cases, but the terminology is not clear. In agreement with previous distinctions, we shall use here the term "weak Heywood" to refer to a 0 uniqueness (or 1 communality) estimate, and "strong Heywood" to refer to a negative uniqueness or greater than 1 communality. However, we also acknowledge that weak Heywood cases or boundary solutions will be very rare in practical applications. We also note that in an orthogonal FA solution, the communality of an item is the sum of the squared factor loadings of this item. So, an "inflated" or overestimated communality estimate implies that some of the factor loadings for this variable would also be inflated/overestimated.

The expected consequences can now be stated as follows: If a ULS-based EFA is fitted to an indefinite **R**, and solutions with different numbers of common factors are fitted until they expand the common factor space, then inadmissible estimates in the form of Heywood cases will necessarily appear. If **R** has substantial negative eigenvalues, Heywood cases might appear even in unidimensional solutions. If it is only slightly indefinite and only solutions with few factors are attempted, they might not appear. However, even in this case some loading estimates are expected to be inflated or overestimated (Cureton & D'Agostino, 1983). These "sequential" results – that is, (a) Heywood cases will appear when solutions with an increasing number of factors are fitted to **R**, and (b) the stronger the indefiniteness of **R**, the sooner they are expected to appear – are the bases of the procedure proposed in this article.

An illustrative, didactic example of these results based on a graphical representation is provided below.

A graphical representation of the problem based on a data example

Let us take a set of six items that are to be responded using a YES/NO format (i.e., they are binary items) and are assumed to measure two common factors. In order to assess this dimensionality hypothesis, a researcher should administer the six items to a sample of a given size, compute the tetrachoric correlation matrix and finally fit a two-factor solution. Table 2 shows **R** in the population and the corresponding 'true' pattern matrix. These are the matrices that the researcher aims to estimate using his/her sample.

It must be noted that **R** in the population is PD, as it should be, with a lowest eigenvalue equal to.516. Figure 1 prints the graphical representation of the six items projected on the two common factors. The blue points represent the position of each item in the space defined by the two factors: the position of each item is defined by the loadings of the item on the two factors. The graph is based on the population loading values

Table 2. Correlation and loading matrices that the researcher needs to estimate from a sample.

	R population matrix					Factor solution: Loading matrix		
	I1	I2	I3	I4	I5	F1	F2	H
I1						.537	.036	.290
I2	.375					.704	.020	.496
I3	.203	.264				.365	.414	.305
I4	.036	.032	.241			.029	.542	.294
I5	.039	.025	.291	.368		.019	.666	.444
I6	.032	.026	.302	.386	.484	.019	.714	.511

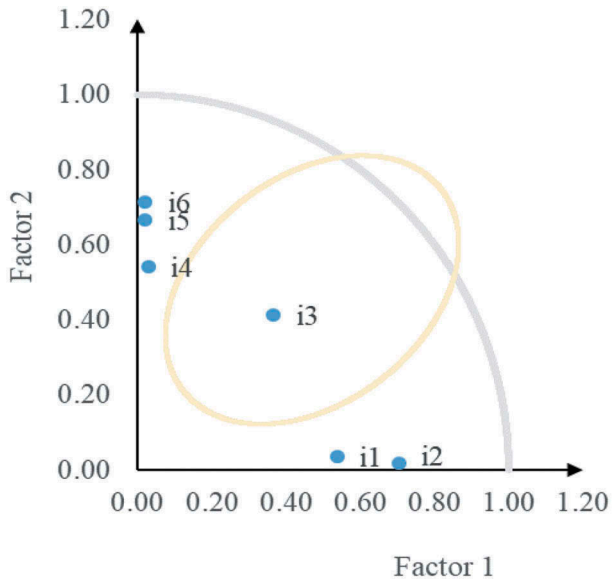

Figure 1. Graphical representation of the two-factor model solution in the population.

obtained from the population **R** matrix, and is the two-factor model that the researcher aims to estimate from a sample. It must be noted that no item should be beyond the gray arch defined by the two factors: any item situated just inside the arch would be a weak Heywood, and outside the arch a strong Heywood (i.e., an item with a communality larger than 1). Table 2 and Figure 1 show that a feasible two-factor model exists for the six items.

So far, the researcher should have no difficulties in estimating the population matrices in Table 2 from a sample and fitting the two-factor model in Figure 1. Now, when a researcher aims to estimate the population loading values using a sample **R** matrix, the better the estimates of the correlation values are, the better the estimates of the loading values themselves will be. Unfortunately, our researcher has limited resources and is only able to recruit a sample of 75 people. With such a small sample, a large amount of sampling error is likely to be present in his/her estimates. To help to visualize the amount of sampling error with this small sample, Figure 1 represents the ellipse of the 90%-confidence interval of the possible estimated positions of one of the items in the analysis: item 3 (labeled *i3* in Figure 1). The same ellipse could be drawn for each one of the six items, but we focus on this single item for the simplicity of the graph.

It can be observed that the true position of item 3 is not in the center of the ellipse: this is because the distribution of the

Table 3. Correlation and loading matrices estimated by the researcher from a small sample ($N = 75$).

	Estimate of **R** from a sample					Loading matrix		
	I1	I2	I3	I4	I5	F1	F2	H
I1						.642	−.067	.416
I2	.497					.738	−.162	.572
I3	.408	.408				.720	.734	1.056
I4	.038	−.079	.139			−.096	.523	.283
I5	−.237	−.115	.524	.507		−.082	.713	.515
I6	−.080	−.301	.642	.542	.519	−.091	.889	.799

estimated loading values is expected not to be symmetrical, but skewed. In addition, the size of the ellipse is in fact large because the sample is very small. Finally, it must be noted that a small area of the ellipse is actually outside the gray arch: this means that some of the estimated positions for item 3 would place it outside the admissible solutions. Whenever this happens, the estimated two-factor model will produce a Heywood case, and the estimated factor solution cannot be interpreted. Table 3 shows the estimates that our unfortunate researcher obtained using his/her small sample.

The first clue of the anomalous situation is that the tetrachoric correlation matrix **R** obtained from the sample is indefinite, with a negative eigenvalue of −0.036. However, an inexpert researcher could miss this detail, and go straight to the piece of the output produced by his/her computing software where the loading matrix is printed. The first impression is actually quite good because the two-factor solution is clear and the estimated loading values are quite large (even larger than the ones in the population that he/she aims to estimate). An experienced researcher will not miss, however, that the communality of item 3 is larger than one. Figure 2 shows where item 3 (labeled as *i3a*) has been placed on the basis of the loading values estimated with the small sample.

The conclusion of the example is that while a feasible two-factor model does exist in the population, the use of the small

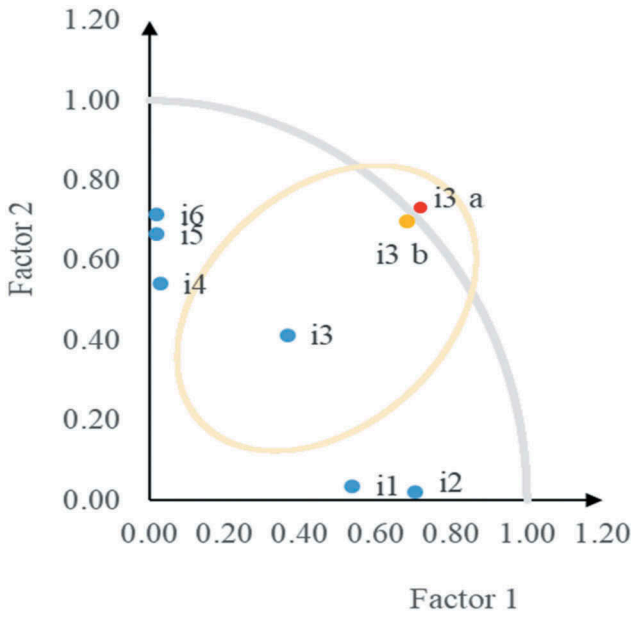

Figure 2. Graphical representation of the two-factor model solution in the population that includes the position of item 3 estimated from a small sample.

sample led to a correlation matrix that was NPD. The consequence was that the estimated loading values are overestimated and a Heywood case appeared. In addition, if the researcher decided to explore if a unidimensional factor model was feasible based on his/her sample **R** matrix, then no Heywood case would be observed in this dataset: the largest estimated loading value would correspond to item 6, with a value of .956. While the unidimensional solution seems to be admissible (as no Heywood case appeared), the estimated loading values would still be overestimated. As an example of this overestimation, consider that when a single factor is extracted in the population, item 6 has a loading value of .676 (instead of the value of .956 estimated from the sample matrix **R**).

Our advice to the researcher would be to substantially increase the size of the sample in order to obtain more reliable estimates of the correlations between items in the population. If this is not possible, then we would advise using a technical approach to correct the sample correlation matrix **R** so that it becomes a PD matrix. The most popular techniques are known as *smoothing techniques*. Figure 2 shows the estimated position of item 3 after the matrix has been smoothed (labeled as *i3b*).

Preventing the problem and existing smoothing solutions

Prevention is the best way to avoid, insofar as this is possible, the problems of NPDs when working with polychoric matrices. Careful selection of the variables (items in our case) and data collection is crucial in any EFA design, and the problem dealt with here might well arise (or be considerably aggravated) from a failure to meet good standards. As described above, redundant or repeated items, non-informative items, and items that all the individuals tend to respond to in the same way must be avoided. And, as for data collection, pairwise deletion should also be avoided whenever possible.

Provided that the recommendations above are met, the problem of NPDs is mainly a problem of sampling error. And, as above, the problem is a specific case of a more general problem. Practitioners are not usually aware that the sampling error of a polychoric correlation is generally far larger than that of a Pearson correlation based on the same data. To see this point, consider that in the extreme case of a tetrachoric correlation, the sampling error is at least 50% larger than that of a Pearson correlation, and can become considerably larger as the correlation becomes smaller and the splits become more extreme (e.g. Guilford & Lyons, 1942). Guilford's recommendation for the size of the sample to be double that used for a Pearson matrix is, in our opinion, a good one.

The consequences of working with small samples are not only that the inter-item matrix is more likely to be NPD. The parameter estimates in EFA (i.e. the factor loadings) are based on the inter-item correlations, and if these correlations are unstable and have large sampling variability, the loading estimates based on them are expected to be even more unstable. A solution of this type cannot be generally trusted and is unlikely to generalize across samples drawn from the same population.

If all the recommendations above are implemented, and the matrix still fails to be PD, then, post-hoc, smoothing corrections that aim to render the correlation matrix at least PSD should be employed. Below, we shall review some of these corrections, and propose a new one. Before doing so, however, it would be good to consider the goals and benefits we expect to obtain from the correction. In our opinion, the most important are:

- To obtain the adequacy measures that require inversion of **R** (Bartlett, KMO).
- To avoid inadmissible (out of their bounds) uniqueness/communality estimates.
- To bring the parameter estimates (communalities, loadings) closer to their parameter values (reduce biases).
- To improve assessment of model-data fit.
- To allow procedures that require **R** to be PD, such as MRFA (ten Berge & Kiers, 1991) to be performed using the corrected matrix. The main advantage of using the MRFA approach is to obtain a proper reduced correlation matrix in which the uniquenesses of the factors not considered in the solution are minimized. This property is very useful for assessing essential dimensionality via the proportion of common variance that can be explained by the retained factors (see e.g. Ferrando & Lorenzo-Seva, 2018).

The basic principle in the smoothing corrections is to change the relative weight of the diagonal elements of the correlation matrix with respect to the non-diagonal elements. Now, a straight approach would be to change all the outside diagonal elements so that they are closer to zero: for example, all the elements could be divided by 100. While most of the NPD correlation matrices would probably be corrected at once with this extreme approach, a large amount of variance would be destroyed in the process (i.e., the new **R** matrix would be close to a unit matrix), and it could not be used to compute any further interesting analyses. So, the challenge of the smoothing procedures is to change the relative weight of the diagonal elements of the correlation matrix while destroying as little variance as possible in the process. We shall now review some of the most popular smoothing procedures available and propose one of our own.

Least-squares smoothing proposed by Knol and ten Berge

Knol and ten Berge (1989) proposed approximating a correlation matrix that is NPD by using the symmetric, unit diagonal, PSD matrix that is the best least-squares fit to **R**. The outcome of the procedure is a smoothed correlation matrix ($\tilde{\mathbf{R}}$) of lower rank that is constrained to be PD. From a practical point of view, when this approach is used, the values of the non-diagonal elements of the correlation matrix increase in comparison to those on the diagonal, which implies that the communality of the set of items is artificially increased.

Linear smoothing adding a ridge proposed by Jöreskog and Sörbom

In this approach a constant value is added to all the diagonal elements of the **R** correlation matrix that is NPD. It was first

proposed for use in the context of factor analysis by Jöreskog and Sörbom (1981). The size of the value added is the one empirically detected in order to make the matrix PD. From a practical point of view, a small value can be added iteratively until all the eigenvalues of the matrix are positive. In this manuscript, whenever we used this approach we added $0.0010/\sqrt{(N)}$. As the sum of the diagonal elements of the smoothed matrix does not equal the number of variables, the smoothed matrix must be considered a variance/covariance matrix. In order to convert it back to a correlation matrix, the smoothed matrix can be standardized as,

$$\widetilde{\mathbf{R}} = \text{diag}\left(\widetilde{\mathbf{S}}\right)^{-\frac{1}{2}}\widetilde{\mathbf{S}}\text{diag}\left(\widetilde{\mathbf{S}}\right)^{-\frac{1}{2}} \quad (1)$$

where diag() refers to the diagonal elements, and $\widetilde{\mathbf{S}}$ is the smoothed variance/covariance matrix. Now $\widetilde{\mathbf{R}}$ is a smoothed correlation matrix that is PD.

Please note that in an extreme situation of a correlation matrix with very large eigenvalues, the final smoothed matrix will be a diagonal matrix with ones on the diagonal and zeros elsewhere (i.e., a unit matrix). This sort of situation means that all the information in the correlation matrix has been destroyed during the smoothing process. In order to quantify the amount of information destroyed, the following index can be computed,

$$v = \frac{\sum_{i\neq j} r_{ij} - \sum_{i\neq j} \tilde{r}_{ij}}{\sum_{i\neq j} r_{ij}} \quad (2)$$

where r_{ij} are the off-diagonal elements of the correlation matrix that is NPD, and $\tilde{r}_{ij}$ are the off-diagonal elements of the smoothed correlation matrix that is PD. A value of v of one would mean that no information has survived the smoothing procedure, while a value close to zero would mean that the amount of information destroyed is minimum. In addition to v, the same information could be computed for each variable in the correlation matrix. This index v_j would be reporting the amount of information destroyed in each variable.

In comparison with the previous approach, the values of the non-diagonal elements of the correlation matrix are decreased in comparison to those on the diagonal matrix, which implies that the communality of the set of items is artificially decreased. This feature is common to all the smoothing methods described from now on.

Non-linear smoothing proposed by Devlin, Gnanadesikan and Kettenring

Devlin, Gnanadesikan, and Kettenring (1975, 1981) proposed subtracting (or adding) a small constant value directly from (or to) the elements outside the diagonal of a correlation matrix that is NPD. The subtraction is applied iteratively until the smoothed matrix becomes PD. The approach is nonlinear because the subtraction (or addition) is only applied to the elements outside the diagonal of the smoothed matrix that are not zero. The subtraction is applied to positive values in the smoothed matrix, while the addition is applied to the negative values. Again, extreme situations could lead to a smoothed matrix that is diagonal and from which all the information has been removed during the smoothing process. The amount of information destroyed during the smoothing process can also be quantified using expression (2).

Straight smoothing proposed by Bentler and Yuan

Bentler and Yuan (2011) observed that the above strategies impacted all the variables in the correlation matrix. To prevent this from happening, they proposed focusing the smoothing procedure only on the problematic variables (i.e., the ones that would potentially produce a Heywood case in the factor solution). Their proposal is to extract all the possible factors in the common factor space and to check which variables have communalities larger than 1 in the factor solution. Once these variables have been detected, the correlation estimates to which they are related are decreased by a low value k, so the smoothed $\widetilde{\mathbf{R}}$ matrix is PD. The decreasing factor is arbitrary and depends on the correlation matrix at hand. In their numerical example, they used a value of $k = .96$ but any other value can be used as long as the information lost in the smoothed $\widetilde{\mathbf{R}}$ matrix is minimum, and the matrix $\widetilde{\mathbf{R}}$ is PD. Whenever we computed this approach, we implemented their method iteratively in order to find an optimal constant value k. In the first iteration, we multiplied the corresponding values of $\mathbf{R}$ by a value of $k = 1–0.0010/\sqrt{(N)}$, and decreased k progressively with the value $0.0010/\sqrt{(N)}$, until the smoothed correlation matrix $\widetilde{\mathbf{R}}$ was PD.

It must be noted that the information lost during the smoothing procedure only has an effect on a (possibly small) number of variables. In the most extreme situations, all the information in the smoothed variables could be lost, which would be equivalent to removing these variables from the analysis. Again, the amount of information lost in $\widetilde{\mathbf{R}}$ can be quantified using expression 2.

As Bentler and Yuan did not explicitly name their approach, here we refer to it as "straight smoothing". This label refers to the fact that the method attacks all the possible annoying variables at once. Debelak and Tran (2013, 2016) concluded that this approach was the best option among the smoothing methods they compared in the context of principal components and parallel analysis.

A new proposal: sweet smoothing

Finally, we propose a new approach that is essentially equivalent to that of Bentler and Yuan, but applied very carefully, so that the amount of lost information in $\widetilde{\mathbf{R}}$ is minimal. Our approach can be summarized with this iterative algorithm:

Step 1. Set the number of factors to be extracted $r = 1$.

Step 2. Extract r factors from $\mathbf{R}$, and check for Heywood cases.

Step 3. If no Heywood cases are observed, increase r in 1 and go to step 2. Otherwise, go to next step.

Step 4. Set the correction value $k = 1–0.0001$.

Step 5. Decrease the correlation values of the variables that showed communalities larger than 1 using the value k, in order to obtain the smoothed correlation matrix $\widetilde{\mathbf{R}}$.

Step 6. Check if the matrix $\widetilde{\mathbf{R}}$ is PD: in this case, the algorithm ends and matrix $\widetilde{\mathbf{R}}$ is the smoothed correlation matrix that removes the minimum information. Otherwise, go to next step.

Step 7. Decrease k with the value .0001.

Step 8. If the value of k is lower than .5 and r is lower than the maximum possible number of factors in the common factor space, then it is considered that too much information is to be removed from the variables at hand: in this case, increase r in 1 and go to step 2. Otherwise, go to step 5.

While all the information in some of the smoothed variables could be lost, the algorithm is expected to find the minimum number of variables that need to be smoothed, and aims to produce the minimum loss of information. In order to achieve it, the smoothing of the NPD matrix is done progressively and very carefully, removing a very low amount of information in each iteration. It could be said that we manipulated the NPD matrix in a loving way with the aim of damaging it as few as possible. This is why we label it *sweet smoothing*.

Comparison between straight smoothing and sweet smoothing

In order to compare the two approaches and to understand whether they actually performed differently, we computed a short simulation study. We produced a set of 7 (dichotomous) items that in the population would be modeled by a two-dimensional factor model. The corresponding pattern matrix in the population is shown in Table 4. The reader can observe that: (a) items 1 and 2 are pure indicators of Factor 1; (b) items 5 and 6 are pure indicators of Factor 2; (c) items 3 and 4 are complex items; and (d) items 7 and 8 are not indicators of any of the two factors. Instead of analyzing the correlation matrix between the items expected in the population, we drew small samples of size $N = 75$, until 1,000 **R** matrices that were NPD (with a negative eigenvalue between −0.10 and −0.15) were obtained. Then we smoothed the matrices using *Straight Smoothing* and *Sweet Smoothing*. We observed that *Straight Smoothing* modified the correlation of an average of 2.96 items, and the correction value k was on average equal to .896. On the other hand, *Sweet Smoothing* modified the correlation of an average of 1.57 variables, and the correction value k was on average equal to .825. The first approach modified the value of more variables, while the second removed a bit more of information from the variables it modified. Table 4 also shows the percentage of times that each item was modified. Both methods quite systematically modified item 4, which was the most complex and had most variance. In addition, *Sweet Smoothing* modified items 7 and 8 (the ones with lowest communality) on few occasions. It may be surprising that *Straight Smoothing* frequently modified the values of these two items: the explanation is that both items are affected by a large amount of sampling error, and their communality is often considerably overestimated when the common factor space is explored.

Table 4. Differences of performance between straight and Sweet Smoothing algorithms.

	Loading matrix in population		Percentage of times that each variable needed to be smoothed	
Items	F1	F2	Straight Smoothing	Sweet Smoothing
1	.8	.0	33%	20%
2	.8	.0	34%	18%
3	.6	.6	28%	16%
4	.7	.7	54%	56%
5	.0	.8	32%	19%
6	.0	.8	35%	18%
7	.1	.1	38%	9%
8	.1	.1	39%	9%

How this different performance affects the factor model estimates in general needs to be explored in a more exhaustive simulation study. Interested readers can obtain the simulation study from the authors by e-mail, or download it from the website http://psico.fcep.urv.cat/utilitats/factor/Documentation.html.

Implementing Sweet Smoothing in FACTOR

The authors' experience suggests that proposals such as the present one are only used in practical applications if they are implemented in user-friendly and easily available software. In this respect, the procedure proposed here has been implemented in the 10.10 version of the program FACTOR (Ferrando & Lorenzo-Seva, 2017). Although researchers can select between various smoothing procedures, the default option is sweet smoothing.

In addition, packages in R can be found that implement different smoothing algorithms. For example, Waller (2019) developed the package *fungible* that includes the function *smoothBY* that correspond to the algorithm that we labeled here as Straight Smoothing (Bentler & Yuan, 2011). As our Sweet Smoothing is closely related Straight Smoothing, it should be easy to implement it base on the available R function.

Discussion and conclusions

Polychoric inter-item correlation matrices that fail to be PD are relatively common in factor analytic applications, particularly in scenarios that use the exploratory version of the model. It is generally the cause of considerable concern among EFA practitioners. It is indeed submitted that the causes of NPDs are well known among methodologists and psychometricians, and also that there are excellent general guidelines such as Wothke's (1993). However, to the best of our knowledge, this is the first didactic approach to the problem that (a) specifically discusses item EFA, and (b) focuses on limitations and improper solutions. We believe that this part of the article will provide useful guidelines for practitioners, both in terms of preventing the problem from appearing and of correcting it when this is the only feasible option. In our experience, practitioners are not generally aware of many of the points discussed here, particularly the sampling error problems associated to the use of polychorics, and the links between NPDs and improper solutions.

Our methodological contribution is that we refine an existing procedure that smooths an indefinite **R** with minimum intrusiveness and destroys the minimum possible amount of information. Sweet smoothing appears to fulfil these purposes: it renders **R** PD, produces minimum biases compared to other smoothing alternatives, and seems to improve goodness-of-fit assessment not only with respect to existing alternatives, but also with respect to the option of leaving **R** untransformed. At the practical level, the proposal has been implemented in

a noncommercial, free and popular program. So, we expect the proposal to be used in applications.

Wothke (1993) criticized standard smoothing schemas for only fixing the data but not diagnosing the sources of NPDs. This is not the case with sweet smoothing. Our proposal allows the practitioner to detect which items are responsible for the NPDs of **R**, and this information is potentially useful in practical applications. In principle, our simulation results suggest that the items implied in the correlations that have large sampling errors, high communality items, and complex items are most prone to causing these problems.

We acknowledge that our proposal has its share of limitations and points that require further research. Thus, more extensive simulations would be of interest, and further empirical studies are needed to ascertain the appropriateness of our proposal in practice. And, as far as the scope is concerned, our article has dealt only with ULS estimation. While the reasons for doing so are, in our opinion, justified, it would be of interest to extend both the discussion and the study of the performance of sweet smoothing to other popular least squares schemas in item FA such as WLS and DWLS. Finally, we have only considered the causes and remedies of NPDs in terms of the structural (calibration) stage of FA. It would be interesting in further studies to assess the impact of NPDs in the scoring stage and whether the use of our proposed correction is able to lead to improved factor score estimates.

Acknowledgments

This project has been made possible by the support of the Ministerio de Economía, Industria y Competitividad, the Agencia Estatal de Investigación (AEI) and the European Regional Development Fund (ERDF) (PSI2017-82307-P).

Funding

This work was supported bythe Ministerio de Economía, Industria y Competitividad, the Agencia Estatal de Investigación (AEI) and the European Regional Development Fund (ERDF) [PSI2017-82307-P].

ORCID

Urbano Lorenzo-Seva http://orcid.org/0000-0001-5369-3099
Pere J. Ferrando http://orcid.org/0000-0002-3133-5466

References

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438. doi:10.1080/10705510903008204

Barendse, M. T., Oort, F. J., & Timmerman, M. E. (2015). Using exploratory factor analysis to determine the dimensionality of discrete responses. *Structural Equation Modeling*, *22*, 87–101. doi:10.1080/10705511.2014.934850

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*, 77–85. doi:10.1111/bmsp.1950.3.issue-2

Bentler, P. M., & Yuan, K. H. (2011). Positive definiteness via off-diagonal scaling of a symmetric indefinite matrix. *Psychometrika*, *76*, 119–123. doi:10.1007/s11336-010-9191-3

Cattell, R. B. (1986). The psychometric properties of tests: Consistency, validity and efficiency. In R. B. Cattell & R. C. Johnson (Eds.), *Functional psychological testing* (pp. 54–78). New York, NY: Brunner/Mazel.

Cureton, E. E., & D'Agostino, R. B. (1983). *Factor analysis: An applied approach*. Hillsdale, NJ: LEA.

Debelak, R., & Tran, U. S. (2013). Principal component analysis of smoothed tetrachoric correlation matrices as a measure of dimensionality. *Educational and Psychological Measurement*, *73*, 63–77. doi:10.1177/0013164412457366

Debelak, R., & Tran, U. S. (2016). Comparing the effects of different smoothing algorithms on the assessment of dimensionality of ordered categorical items with parallel analysis. *PLoS One*, *11*, e0148143. doi:10.1371/journal.pone.0148143

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika*, *62*, 531–545. doi:10.2307/2335508

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1981). Robust estimation of dispersion matrices and principal components. *Journal of the American Statistical Association*, *76*, 354–362. doi:10.1080/01621459.1981.10477654

Ferrando, P. J., & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica*, *21*, 301–323.

Ferrando, P. J., & Lorenzo-Seva, U. (2014). Exploratory item factor analysis: Additional considerations. *Anales de psicologia*, *30*, 1170–1175.

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, *29*, 236–240. doi:10.1007/978-4-431-68225-7_2

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. doi:10.1177/0013164417719308

Foldnes, N., & Grønneberg, S. (2019). On identification and non-normal simulation in ordinal covariance and item response models. *Psychometrika*, *84*, 1000–1017. doi:10.1007/s11336-019-09688-z

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*, *16*, 625–641. doi:10.1080/10705510903203573

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.

Guilford, J. P., & Lyons, T. C. (1942). On determining the reliability and significance of a tetrachoric coefficient of correlation. *Psychometrika*, *7*, 243–249. doi:10.1007/BF02288627

Harman, H. H., & Jones, W. H. (1966). Factor analysis by minimizing residuals (Minres). *Psychometrika*, *31*, 351–369. doi:10.1007/bf02289468

Heywood, H. B. (1931). On finite sequences of real numbers. *Proceedings of the Royal Society of London. Series A*, *134*, 486–501.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL 5: Analysis of linear structural relationships by maximum likelihood and least squares methods;[user's guide]*. Uppsala, Sweden: University of Uppsala.

Kaiser, H. F. (1970). A second generation of Little Jiffy. *Psychometrika*, *35*, 401–415. doi:10.1007/bf02291817

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, *34*, 111–117. doi:10.1177/001316447403400115

Kim, S., Lu, Z., & Cohen, A. S. (2020). Reliability for tests with items having different numbers of ordered categories. *Applied Psychological Measurement*.

Knol, D. L., & Berger, M. P. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research*, *26*, 457–477. doi:10.1207/s15327906mbr2603_5

Knol, D. L., & ten Berge, J. M. (1989). Least-squares approximation of an improper correlation matrix by a proper one. *Psychometrika*, *54*, 53–61. doi:10.1007/BF02294448

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. doi:10.1080/00273171.2012.658340

Lee, S. Y., & Poon, W. Y. (1987). Two-step estimation of multivariate polychoric correlation. *Communications in Statistics-Theory and Methods*, *16*, 307–320. doi:10.1080/03610928708829368

Lorenzo-Seva, U., & Van Ginkel, J. R. (2016). Multiple imputation of missing values in exploratory factor analysis of multidimensional scales: Estimating latent trait scores. *Anales de Psicología/Annals of Psychology*, *32*, 596–608. doi:10.6018/analesps.32.2.215161

McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: LEA.

McDonald, R. P., & Ahlawat, K. S. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology*, *27*, 82–99. doi:10.1111/bmsp.1974.27.issue-1

Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics*, *11*, 3–31. doi:10.3102/10769986011001003

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*, 115–132. doi:10.1007/BF02294210

Muthén, B. (1993). Goodness of fit with categorical and other nonnormal variables. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (Vol. 154, pp. 205–234). Newbury Park, CA: Sage. doi:10.1093/sf/73.3.1161.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*, 443–460. doi:10.1007/BF02296207

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287–297. doi:10.1037//1040-3590.12.3.287

ten Berge, J. M. F. (1998). Some recent developments in factor analysis and the search for proper communalities. In A. Rizzi, M. Vichi, H. H. Bock (Eds.), *Advances in data science and classification* (pp. 325–334). Berlin, Germany: Springer.

Ten Berge, J. M. F., & Kiers, H. A. L. (1991). A numerical approach to the exact and the approximate minimum rank of a covariance matrix. *Psychometrika*, *56*, 309–315. doi:10.1007/bf02294464

Waller, N. G. (2019). Fungible: Psychometric functions from the Waller Lab. *R Package Version*, *1*, 93.

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods*, *12*, 58. doi:10.1037/1082-989X.12.1.58

Wothke, W. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Nonpositive definite matrices in structural modeling* (pp. 256–293). Newbury Park, CA: Sage. doi:10.1093/sf/73.3.1161.

Yuan, K. H., Wu, R., & Bentler, P. M. (2011). Ridge structural equation modelling with correlation matrices for ordinal and continuous data. *British Journal of Mathematical and Statistical Psychology*, *64*, 107–133. doi:10.1348/000711010X497442

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. doi:10.2333/bhmk.33.61