24 Introduction to Exploratory Factor Analysis: An Applied Approach

Martin Sellbom and David Goretzko

Abstract

This chapter provides an overview of exploratory factor analysis (EFA) from an applied perspective. We start with a discussion of general issues and applications, including definitions of EFA and the underlying common factors model. We briefly cover history and general applications. The most substantive part of the chapter focuses on six steps of EFA. More specifically, we consider variable (or indicator) selection (Step 1), computing the variance–covariance matrix (Step 2), factor-extraction methods (Step 3), factor-retention procedures (Step 4), factor-rotation methods (Step 5), and interpretation (Step 6). We include a data analysis example throughout (with example code for R), with full details in an online supplement. We hope the chapter will provide helpful guidance to applied researchers in the social and behavioral sciences.

Keywords: Exploratory Factor Analysis; Factor Analysis; Internal Structure; Measurement Modeling; Latent Variable Modeling

Introduction

Exploratory factor analysis (EFA) has been an incredibly popular statistical technique in the social and behavioral sciences for over a century (e.g., Goretzko et al., 2021). In the first author's field of psychological assessment, for instance, EFA has been used to elaborate on the internal structure of psychological tests since its inception, and it remains popular today. Open any issue of *Psychological Assessment*, *Journal of Personality Assessment*, or *Organizational Research Methods*, to mention just a few, and you are bound to see articles that used EFA as a method to develop and/or evaluate operationalizations of various constructs. Although popular in psychology, the importance of EFA has been identified in many other areas of the social and behavioral sciences, such as sociology (e.g., Kirkegaard, 2016), education (e.g., Beavers et al., 2013), organizational research (e.g., Conway et al., 2003), and communication science (e.g., Park et al., 2002).

Precursors to contemporary EFA have been available almost as soon as correlation matrices could be calculated (see Mulaik, 2010 for a review). As Sir Francis Galton and Karl Pearson worked on mathematical models of correlation, that would

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

ultimately yield the still-popular Pearson product moment correlation coefficient (Pearson, 1909), other scholars used these methods to calculate intercorrelation matrices to evaluate higher-order indices for interrelated variables. Charles Spearman, for example, built a higher-order model of intelligence (Spearman, 1904). Spearman's work was actually more reminiscent of the bifactor model approach, which has become quite popular recently, than EFA (see e.g., Sellbom & Tellegen, 2019). It was subsequent scholars who ultimately advocated for the EFA principles and methods that are frequently used today (e.g., Cattell, 1943; Thurstone, 1938).

This chapter emphasizes EFA. There are several forms of data reduction techniques, such as principal components analysis (PCA) and image factor analysis, which make different assumptions about the variances in the variables (i.e., indicators) being analyzed. We also focus specifically on latent variable models and do not provide coverage of other related methods of evaluating structure of variables, such as network analysis (e.g., exploratory graph analysis) or person-centered cluster analytical approaches. Furthermore, we do not cover confirmatory factor analysis (CFA) – a special case of structural equation modeling (SEM, see Chapter 25 in this volume) – though many issues pertaining to indicator selection and estimators apply to CFA as well. The "SEM Steps and Reporting Standards" in Chapter 25 of this volume apply to specifying, estimating, and evaluating a CFA model; Figure 25.2 and its associated narrative discussion provides a good example in that chapter. We further note that this chapter is mostly applied in nature, meaning that we do not take a mathematical approach to explaining the conceptual and practical foundations of EFA. Indeed, any reader interested in such foundations is referred to Mulaik's (2010) excellent book on this topic.

Definitions and Contrasting from Other Methods

Readers may rightfully wonder how EFA is different from some of these other data reduction methods just mentioned. We will contrast EFA from the two most common other alternatives (PCA and CFA). First, PCA and EFA are often confused as they are both similar forms of data reduction, and similar steps are applied in selecting the optimal structure – rotating solutions to simple structure, theoretical evaluation of competing structures, etc. However, it is important to note that, even if PCA and EFA methods often yield similar solutions, they are based on different assumptions about the underlying variances in the variables being analyzed.

EFA is based on the *common factor model* (e.g., Thurstone, 1947), which assumes each variable in a set of observed or measured variables (i.e., *indicators*) is a linear function of one or more unobserved (i.e., *latent*) factors as well as a residual factor unique to each variable (i.e., a *unique variance* component). Each latent variable is estimated through the variance common across the set of indicators (hence, common factors) and, specifically, that is being predicted by the latent variable. In other words, the underlying reason (or "cause") for a particular value on any observed indicator is the level of the underlying latent construct. The common factor model

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

also considers residual factors that represent the unique variance of each indicator when the common variance has been accounted for; this unique variance is a combination of both systematic (or reliable) influences that are unrelated to the latent variable(s) as well as unsystematic (or unreliable) variances.

PCA, on the other hand, is not based on the common factor model as it does not parcel out shared and unique variances (Mulaik, 2010). Rather, PCA is a more simplistic procedure that attempts to maximize the amount of variance for which can be accounted in the indicators rather than making assumptions about causation. Brown (2014) points out that some scholars nonetheless argue that PCA might be advantageous to EFA because it is more simplistic mathematically, is less prone to problematic solutions, is not hampered by factor indeterminacy (i.e., component scores can be calculated more easily than factor scores), and PCA and EFA often yield similar results. However, as also noted by Brown (2014), other scholars (e.g., Fabrigar et al., 1999; Floyd & Widaman, 1995; see also Schmitt, 2011) have generally refuted these arguments because solutions are indeed dissimilar under various conditions (e.g., few indicators per factor and small communalities – the amount of variance accounted for in an indicator by all factors); more generally, analyses should be applied based on the underlying theoretical assumptions made about associations among variables. Moreover, because both EFA and CFA are based on the common factor model, EFA results are more likely to be supported by subsequent CFA in other samples (Floyd & Widaman, 1995; Schmitt, 2011).

The primary difference between EFA and CFA are in the names. Exploratory methods make no a priori assumptions about structure and are best suited for contexts in which the underlying structure of a set of variables is unknown. Confirmatory analyses, on the other hand, are explicitly testing one or several competing theoretical structures that are indeed known. Both are based on the common factor model, but one important difference in CFA is the reliance on the independent clusters model. In other words, a standard CFA model typically assumes one cause (i.e., latent factor) per indicator. EFA, on the other hand, makes no such assumption and estimates all latent factors as predictors for all indicators in the model; instead, it uses rotation methods (defined in a later section) to view a particular solution from a simple structure perspective. Finally, unlike EFA, the global evaluation of CFA models is largely based on the degree to which the specified model is consistent with the observed data (i.e., model fit) and statistical comparison to other theoretically plausible models. See Chapter 25 for detailed coverage of these issues in the broader SEM context.

General Applications

EFA can be useful in any context in which higher-order explanations of intercorrelations among a set of variables can be beneficial. Indeed, it has been used to articulate the structure of major psychological constructs, including intelligence (e.g., Thurstone & Thurstone, 1941) and personality (e.g., Cattell, 1945). A particularly common application of EFA is the examination of the internal

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

structure of psychological test items (Brown, 2014), especially when no clear a priori theoretical structure exists. For instance, imagine a researcher has developed a new self-report questionnaire for assessing educational learning strategy – a multidimensional construct composed of multiple abilities (e.g., Berger & Karabenick, 2016). Therefore, the researcher, who has developed 25 test items to measure important features of learning strategies, conducts an EFA to determine the underlying structure of the test.

As another recently published example, Jokiniemi et al. (2021) developed a clinical nurse specialist core competency scale in a sample of nurses from a variety of Nordic countries. Because the underlying structure of the scale was unknown, they subjected 50 items to an EFA and decided on a four-factor structure. The respective loadings of items on the factors indicated four competency spheres of patient, nursing, organization, and scholarship, which formed a final scale.

Steps in Conducting EFA

In this section, we articulate the various steps in the applied use of EFA. For each, we discuss important issues that EFA users should consider and the general empirical literature to guide decisions. We also provide exemplary *R* code for practitioners to illustrate how to conduct the steps 2–5. A more detailed data example can be found in our Open Science Framework (OSF) repository (https://osf.io/srv8e/). The data were provided by Schödel et al. (2018) and consist of 312 observations of 60 extraversion items (four-point Likert scale) from the Big Five Structure Inventory (BFSI; Arendasy 2009); the BFSI measures the five-factor model of personality – a popular personality perspective in psychology. Extraversion is conceptualized as a broad individual differences trait domain with multiple specific trait facets (e.g., gregariousness, warmth, assertiveness), and thus, an extraversion item pool should be multifactorial. Throughout this chapter, we use these data to walk the reader through the basic steps of EFA; for a more detailed depiction, we encourage the reader to study the supplemental material in our OSF repository.

Step 1: Variable Selection

Every EFA begins with variable (indicator) selection, which is dictated by the purpose of the analysis. It is important to keep in mind that the results of EFA are completely bound by the variables included; there is no magic that will reveal some broad truth. Thus, EFA should not be used to articulate theory but, rather, thinking carefully about variable selection should precede the analysis. Of course, for certain applications, the researchers are bound to a particular variable pool. For instance, the evaluation of the internal structure of a psychological test is directly linked to the available items on that test.

It is very important that EFA users pay close attention to the nature of their indicators (e.g., scaling, distribution, degree of unidimensionality), as these properties have important implications for the selection of factor-extraction methods

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

discussed later. Indeed, considerations for scaling/distribution of indicators will be covered under Step 3. In this first step, we consider some issues concerning the nature of indicators to which we believe EFA users should pay particular attention.

In an excellent article on factor analysis, Schmitt (2011) argues that indicators that are highly skewed can, for that reason alone, be highly correlated and indicative of an artefactual factor (see also Sellbom & Tellegen, 2019). This can be illustrated through a simple example. Let's consider two scales from the Minnesota Multiphasic Personality Inventory-2 – Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008). Substance Abuse (SUB; seven items) and Anxiety (AXY; five items) measure two theoretically distinct constructs. We subjected its seven binary (true/false) items to an EFA using a robust weighted least-squares estimator in the modeling software Mplus 8.4 across two separate samples. The first sample consisted of 895 individuals from a community mental health center (Graham et al., 1999). The median endorsement frequency for these items was 27.5% (range: 9.3% to 51.6%). An EFA supported two factors (based on parallel analysis; see Step 4 later), with all seven SUB items loading on the first factor (median = 0.70; range: 0.45 to 0.87) and all AXY items loading on a second factor (median = 0.69; range: 0.48 to 0.79).

The second sample consisted of 336 individuals who had been administered the MMPI-2-RF as part of a pre-employment evaluation for a law enforcement position (Detrick et al., 2016) – a context in which endorsing substance abuse and anxiety symptoms is unlikely to occur due to either good psychological adjustment in such individuals or significant under-reporting. Indeed, the median endorsement frequency of these 12 items was 0.8% (range: 0.3–18.1%). The EFA suggested a clear one-factor solution with all but one of the items loading meaningfully on this factor (median = 0.84, range: 0.37–0.95); the only item that failed to reach a meaningful loading (0.28) was associated with the highest response endorsement (18.1%), and the item with the lowest meaningful loading (0.37) was associated with the second highest response rate (14.8%). Forcing a two-factor solution would result in an improper solution, with the single item left out of the one-factor model forming its own factor with a loading of 1.06. Thus, this example clearly demonstrates the effect that similar and extreme item skew can lead to theoretically inconsistent and artefactual factor solutions.

Indicator parceling is another important issue that EFA users should consider. Parceling refers to adding multiple indicators into a smaller set of aggregates to reduce model complexity. There is debate in the field about whether parceling is appropriate in factor analysis (e.g., Bandalos, 2008; Little et al., 2013; Marsh et al., 2013). Proponents for parceling argue for parcels being more reliable and distributionally sound indicators than the original ones, as well as the benefits of reducing model complexity and increase in statistical power (Little et al., 2013). Opponents, however, argue that parceling can mask potential problems associated with individual indicators, as poor indicator performance could be indicative of problematic content contributing construct-irrelevant variance (e.g., Bandalos, 2008; Marsh et al., 2013). We do not take a strong stance other than to say that if the indicators subjected to EFA meet the general goal of the analysis, and they are sufficiently unidimensional/reliable, indicator parceling is generally appropriate.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Step 2: Compute the Variance–Covariance Matrix

A factor analysis is a test of a variance–covariance matrix (or, in standardized terms, a correlation matrix). Every statistical software will calculate a variance–covariance matrix and subject this matrix to an EFA. In most cases, EFA users do not need to do this themselves. The statistical program will automatically do the conversion based on the instructions received. However, it is also possible for the applied user to calculate a variance–covariance matrix and directly subject this matrix to EFA to make adjustments to the correlations in a manner not possible when using raw indicator data (e.g., dis-attenuating correlations for range restriction or converting a correlation matrix to fit distributional assumptions); a full discussion of these issues is beyond the scope of our chapter. Here, we focus on some important assumptions about the variance–covariance matrix for EFA.

Prior to an EFA being conducted, the user should check whether the variance–covariance matrix meets the assumptions necessary to be subjected to the analysis. There are two common tests. First, the Kaiser–Meyer–Olkin (KMO) index of sampling adequacy directly allows for the examination of whether the variables to be included in the EFA are appropriate for factor analysis as a set. If the common variance across the indicators is too small, it is not meaningful to conduct an EFA. Hence, the KMO – a measure of the proportion of common variance across all indicators – states how well the set of indicators is suited for EFA. Values that are close to 1.0 are preferable (see Kaiser & Rice, 1974 for more guidance on KMO interpretation). The R package psych provides a function to calculate the KMO measure: psych::KMO(efa_data). For our example data on 60 extraversion items, the KMO measure is 0.92 and suggests that the data can be subjected to an EFA.

Another method to determine the suitability of an indicator set for EFA is the Bartlett's test of sphericity that can be used to ensure that the correlation matrix is not an identity matrix. Specifically, if the indicators are not related to one another, no informative structure can be detected. A chi-square test is calculated to test the null hypothesis that the indicators are orthogonal; a significant test, therefore, is evidence that they are not and that the underlying matrix has sufficient covariation to be suitable for EFA. In R, users can apply the cortest.bartlett function of the psych package – psych::cortest.bartlett(cor(efa_data), n = nrow(efa_data)), with cor(efa_data) calculating the correlation matrix and nrow(efa_data) returning the sample size. Bartlett's test of sphericity rejects the null hypothesis in our data example, so we deem our data suitable for an EFA.

Step 3: Factor Extraction

When conducting an EFA, researchers can choose between several factor-extraction methods (i.e., estimation methods). This decision can have substantial influence on the results – especially on the estimated factor loadings (Beauducel, 2001; De Winter & Dodou, 2012). Although it has often been argued that PCA relies on different model assumptions (see earlier discussion on this topic) and, therefore, should not be

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

treated as an alternative extraction method (e.g., Farbrigar et al., 1999) as it may yield biased factor loadings (Widaman, 1993), it is not always clear which other method is preferable (based on the common factor model).

Principal Axis Factoring

Historically, principal axis factoring (PAF) has been the most popular extraction method and remains the preferred method for many researchers who use EFA (Conway & Huffcutt, 2003; Goretzko et al., 2021; Henson & Roberts, 2006; Howard, 2016). Its popularity may be explained by its conceptual resemblance to PCA and its importance in the early days of EFA usage (e.g., Holzinger, 1946). The basic idea of PAF is to adjust the PCA approach to the common factor model to account for measurement error and to consider unique variance components. Instead of decomposing the correlation (or variance–covariance) matrix to find principal components, it works with a so-called "reduced" correlation matrix that contains communality estimates on the diagonal. Hence, the principal axes (factors) obtained from this procedure are not able to explain all the variance of the manifest variables but only shared variances according to the common factor model (i.e., the variance components that are explained by the underlying latent factors that represent, for example, psychological constructs).

The initial estimation of the communalities used to generate the reduced correlation matrix is usually based on the squared multiple correlations (SMCs) among the indicators – the default in statistic programs (e.g., SPSS or R when using the psych library: psych::fa(efa_data, nfactors = 6, fm = "pa", SMC = TRUE)). After initially "guessing" the communalities, an eigenvalue decomposition is performed on the reduced correlation matrix (similar to PCA), and the resulting factor pattern can then be used to re-estimate the communalities. This iterative procedure is continued until a convergence criterion is fulfilled.

(Weighted) Least-Squares Approaches

With this procedure, PAF implicitly aims at finding the factor loadings that minimize the squared deviation between the diagonals of the reduced correlation matrix (which contains the communality estimates) and the reproduced correlation matrix (which consists of the model-implied correlations calculated from the estimated factor loadings; see Jöreskog et al., 2016 for more details). While PAF uses the iterative procedure to estimate loadings and communalities via eigenvalue decomposition, Harman and Jones's (1966) minimizing residuals factor analysis (Minres) determines the factor loadings by reproducing the off-diagonal elements of the correlation matrix as closely as possible (thus circumventing the problem of unique variance elements). The authors demonstrate that Minres and PAF result in the same factor solution if the communalities estimated by Minres are used for PAF (Harman & Jones, 1966).

As Minres minimizes the squared distance between the off-diagonal elements of the correlation matrix and the respective correlations implied by the factor model (i.e., the squared residuals), it provides the solution to the unweighted least-squares-fit

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

function (Jöreskog et al., 2016). Accordingly, Minres yields equivalent results as an ordinary least-squares method and can be seen as a representative of diverse least-squares approaches, which all estimate the model parameters by optimizing a fitting function that compares the actual correlation matrix with a model-implied matrix. The different estimation methods – namely generalized least-squares, unweighted least-squares, and maximum-likelihood (ML) methods – can be formulated as weighted least-squares (WLS) approaches (Browne, 1977) and enable the researchers to test the goodness of fit by calculating common model fit indices (e.g., the RMSEA). Robust WLS or diagonally WLS methods have been developed in the context of SEM research to address the problem of biased standard error estimation in WLS approaches (for an overview, see DiStefano & Morgan, 2014). These approaches, including, for example, WLSMV (weighted least-squares mean and variance adjusted) methods, can also be used in EFA. However, robust WLS is often not selected as an extraction method, probably because EFA is usually used for exploring the data and not for model testing (compared to CFA) and, hence, proper standard error estimation is rarely considered by its users.

One of the advantages of these least-squares approaches is that they do not carry distributional assumptions about the indicators and are, therefore, applicable to all types of variables. As described later, research has indicated that these approaches can be particularly useful for ordered categorical or binary variables (e.g., Rhemtulla et al., 2012). However, a WLS analysis usually needs (slightly) larger sample sizes (e.g., Li, 2016; Rhemtulla et al., 2012) and is not the preferred method when using normally distributed indicators.

Maximum-Likelihood Estimation

While general least-squares approaches come without distributional assumptions, ML estimation puts a stronger focus on the data-generating process and is, therefore, arguably a more sophisticated approach to factor analysis – especially since it treats the unique variances as formal model parameters that have to be also estimated (Everitt & Hothorn, 2011). Usually, multivariate normality is assumed but, theoretically, other distributional assumptions can also be made for ML estimation (e.g., Wedel & Kamakura, 2001).

For ML estimation, a fitting function that is closely related to the likelihood function is minimized with respect to the loading parameters as well as the unique variances. The likelihood function indicates how plausible specific parameter values are given the observed data; in this case, that means the plausibility of specific factor loadings and unique variances. Accordingly, the aim of this estimator is to maximize the likelihood that the final set of parameters map onto the observed data (hence, "maximum likelihood"). Moreover, even though the ML estimation of parameters is fairly robust against violations of the normality assumption (e.g., Jöreskog et al., 2016), standard errors and respective significance tests can deteriorate when the actual data-generating process differs from the assumed one. Therefore, several adjustments for robust ML estimation have been developed (e.g., Yuan & Bentler, 1998).

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Comparison of Factor-Extraction Methods

Selecting one of these estimation methods for an EFA can be challenging as their precision and stability vary across different data conditions. PAF is sometimes favored as it produces fewer Heywood cases (i.e., cases in which unique variances are estimated to be negative or correlations estimated to be greater than one) compared to ML estimation (De Winter & Dodou, 2012). PAF also does not require multivariate normality underlying the indicators. However, the initial communality estimates can heavily influence the outcome of PAF; using the complete variances as communality estimates often yields inflated parameter estimates, while using the SMC approach may cause negative eigenvalues (Gorsuch, 1983). Furthermore, PAF does not allow for a direct replication with CFA (for which the default estimator is typically ML) and does not provide fit indices to evaluate model fit. Hence, several authors advocate not to rely on PAF (Conway & Huffcutt 2003; Fabrigar et al., 1999; Goretzko et al., 2021) but rather to use ML estimation especially when multivariate normality can be assumed. Because EFA results should typically be replicated and validated using CFA on a new sample, a likelihood-based estimation procedure seems to be the most suitable.

When the multivariate normality assumption is violated (e.g., when data are based on indicators with few categories), WLS parameter estimation based on polychoric correlations may be an appropriate alternative to ML estimation (Barendse et al., 2015; Schmitt, 2011); it also can be used in CFA and, hence, for direct cross-validation. EFA users should examine their data carefully and evaluate whether a normality assumption holds. Regardless, when ordinal indicators with fewer than five categories are used, WLS is preferred over ML estimation – particularly robust weighted least squares or unweighted least squares (Beauducel & Herzberg, 2006; Goretzko et al., 2021; Li, 2016; Rhemtulla et al., 2012). The "fa" function of the psych package offers numerous estimation methods. Users can select the preferred method by setting the argument "fm", for example, when performing WLS estimation: psych::fa(efa_data, nfactors = 6, fm = "wls"). Since our example data set consists of four-point Likert items that have to be considered as ordinal variables (see also Beauducel & Herzberg, 2006), we decided to rely on WLS estimation (statistical tests and graphical inspection also suggest that multivariate normality is questionable for our data, see https://osf.io/srv8e/).

Step 4: Factor Retention

Selecting the optimal number of factors to retain constitutes a key decision in EFA. Before estimating the loadings and unique variances, the researcher needs to determine the dimensionality (or number of latent factors). Although theoretical considerations should also be taken into account in this decision-making process, this number is primarily inferred from objective data. Over the years, several factor-retention criteria have been developed to estimate the number of latent factors underlying the correlation matrix of the manifest indicators.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Eigenvalues: Kaiser–Guttman and Empirical Kaiser Criterion

Eigenvalues are central characteristics of a matrix that, in the case of correlation matrices, indicate how much variance in the manifest variables can be explained by the respective eigenvectors (i.e., the principal components in PCA). Therefore, eigenvalues of the correlation matrix (or the reduced correlation matrix in PAF) play a central role in determining the number of factors to retain in EFA. The well-known Kaiser–Guttman rule (Kaiser, 1960), often referred to as eigenvalue-greater-one-rule, suggests retaining as many factors as there are eigenvalues greater than one. At the population level, the correlation matrix under the null model (no underlying factors) is simply an identity matrix and all eigenvalues are one. Accordingly, the rationale of the Kaiser–Guttman rule is that an underlying factor should explain more variance than a single variable and should have a corresponding eigenvalue greater than one.

As Breaken and van Assen (2017) explain, this idea may be reasonable on a population level, but is flawed on a sample level due to sampling error. Therefore, the authors developed a new version of this rule – the empirical Kaiser criterion (EKC) – that takes into account the sample size as well as the size of previous eigenvalues when calculating reference eigenvalues that are compared with the empirical eigenvalues (e.g., the second reference eigenvalue is adjusted to account for a very large first eigenvalue corresponding to a dominant first factor). In other words, EKC provides different reference values for each observed eigenvalue (instead of comparing all eigenvalues with the fixed value of one) and promises to be less prone to sampling error. It suggests retaining factors whose eigenvalues are greater than the calculated reference eigenvalues, considering the sample size, the number of indicators, and all previous eigenvalues. Braeken and van Assen (2017) differentiate the restricted EKC (where the reference eigenvalues are at least one) and an unrestricted version with reference eigenvalues that can be even smaller than one.

Scree Test

Another popular method of determining the number of factors is the scree test (Cattell, 1966) and it is also based on the empirical eigenvalues. The idea behind this method is to plot the eigenvalues in a descending order and to determine an "elbow" in this plot, where the change from one eigenvalue to the subsequent eigenvalue is considerably smaller than the difference between the two prior eigenvalues. The assumption is that all factors corresponding to the eigenvalues before this "elbow" can explain substantial amounts of variance while all factors from this position and onwards are insufficient for this purpose. Ultimately, the visual inspection and interpretation of this scree plot is quite subjective.

Parallel Analysis and Comparison Data

The improvement in computational resources have fostered the applicability of simulation-based factor-retention approaches. Parallel analysis (PA; first implemented by

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Horn, 1965) is the best-known factor-retention approach that uses simulated data for comparison. The basic premise of PA is to generate reference values for the empirical eigenvalues based on several simulated data sets of the same size and number of indicators as the empirical data set. After simulating *B* data sets based on the null model (i.e., no underlying latent factors), the mean of the *B* first eigenvalues is compared to the first empirical eigenvalue, the mean of the *B* second eigenvalues is compared to the second empirical eigenvalue, and so on. PA suggests retaining factors as long as the empirical eigenvalue is greater than the reference eigenvalue. Instead of using the mean to aggregate the respective eigenvalues of the *B* data sets, arbitrary percentiles of the eigenvalue distribution can be taken as the reference value (often the 95% percentile). There are also implementations of PA that are based on the eigenvalues of the reduced correlation matrix (see also the comparison of PCA and EFA) and PA varieties using bootstrapped instead of simulated data. Lim and Jahng (2019) provide a more detailed overview of the different versions of PA and their performance under various data conditions.

Ruscio and Roche (2012) developed the comparison data (CD) approach that combines the simulation of comparison data sets (similar to PA) with the model-testing perspective of CFA (see also the section on model fit indices below). Contrary to PA, the simulated data sets do not represent a null model but are based on different factor models while also reflecting the marginal distributions of the indicators. For each factor solution and each comparison data set, the root-mean-squared error (RMSE) between the empirical eigenvalues and the respective comparison eigenvalues is calculated. That is, if *B* comparative data sets are simulated per factor solution, *B* RMSE values per number of factors are obtained. The CD method then tests whether the RMSE values of a two-factor solution are, on average, significantly smaller than those of a one-factor solution. Mann–Whitney U tests are conducted with subsequent numbers until no "significant" improvement is indicated. To avoid underfactoring, the authors suggest an alpha level of 0.30 as a threshold for significance.

Minimum Average Partial Test

Velicer (1976) developed the minimum average partial (MAP) test that aims at determining the number of components to retain in PCA based on averaged, squared partial correlations of the indicators. Although it was designed for PCA, the MAP test is frequently used in the context of EFA (Goretzko et al., 2021). The basic premise is to determine the number of components for which the squared correlations of the indicators are minimal, on average, after the common variance explained by the principal components is controlled for ("partialed out").

Hull Method

The hull method by Lorenzo-Seva et al. (2011) consists of three major steps. First, a set of factor solutions is selected for which a model fit index is calculated (the authors suggest using the comparative fit index). Then, the fit index is plotted against

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

the corresponding degrees of freedom for each factor solution. Subsequently, an elbow in the upper boundary of the convex hull of the plotted points is detected to determine at which point increasing the number of factors does not substantially improve upon the model fit. Unlike the scree test, the position of the elbow in the upper hull can be calculated, making this a less subjective approach. Due to its model-comparison perspective, the hull method necessitates the use of ML or least squares estimators to calculate model fit indices.

Sequential Chi-Square Tests

When using ML EFA, it is possible to test whether a specific number of factors k is sufficient to explain the common variance of an indicator set. If that is the case, a test statistic proportional to the ML fitting function is approximately chi-square distributed (e.g., Everitt & Hothorn, 2011) and the null hypothesis that k factors are sufficient can be tested. This procedure is repeated with subsequent numbers of factors ($k = 1, 2, 3, \ldots$) until the null hypothesis holds.

Fit Indices and Information Criteria

There are also authors who view factor retention as a model selection problem (e.g., Preacher et al., 2013) and, therefore, rely on relative and absolute measures of model fit. When likelihood-based EFA is conducted, information criteria, such as the Akaike information criterion (Akaike, 1987) or the Bayesian information criterion (Schwarz, 1978), can be used to determine which number of factors represents the empirical relations more accurately. As an alternative to information criteria, fit indices known from model testing in the context of CFA (or structural equation modeling; see Chapter 25 in this volume) can also be used to compare different factor solutions with each other (see Preacher et al., 2013 for more details). Some recent scholars have published simulation data that question the utility of model fit indices for the purposes of factor retention, however (Auerswald & Moshagen, 2019; Montoya & Edwards, 2021).

Factor Forest

Recently, a new simulation- and machine learning-based approach for factor retention has been developed by Goretzko and Bühner (2020). The basic idea of the factor forest is to simulate data under all important data conditions of an application context (i.e., considering common sample sizes, realistic ranges for the number of latent factors and the number of manifest indicators, common loading patterns and communalities, etc.) and then to extract specific data characteristics for each simulated data set (e.g., eigenvalues and matrix norms of the correlation matrix). These data characteristics, and the known number of latent factors (the true dimensionality is known since the data are simulated), are then treated as input (independent variables) and target variables (dependent variable or criterion) of a machine learning model that "learns" how the data characteristics and the number of factors are interlinked.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

The trained model is then able to predict the number of factors given the observed data characteristics of an empirical data set. As this procedure is computationally very costly, Goretzko and Bühner (2020) provide a pre-trained model that was trained on nearly 500,000 data sets based on multivariate normality and between one and eight latent factors; the trained model and the analysis scripts can be retrieved from an OSF repository – https://osf.io/mvrau/ or from our repository with a simplified R script – https://osf.io/srv8e/).

Comparison of Factor-Retention Criteria

Although the Kaiser–Guttman rule, the scree test, and PA are the most popular methods to determine the number of factors (Goretzko et al., 2021), simulation studies suggest that only the latter provides comparably good estimates, and the other methods are often not able to retain the correct number of factors (Auerswald & Moshagen, 2019; Fabrigar et al., 1999; Goretzko et al., 2021; Schmitt et al., 2018). Therefore, PA is seen as the "gold standard" of factor retention (e.g., Braeken & van Assen, 2017; Schmitt et al., 2018); this may also be explained by its relative robustness against distributional assumptions (Dinno, 2009). However, some modern alternatives (e.g., the CD method, EKC, or hull method) have shown advantages over PA in some data conditions (Braeken & van Assen, 2017; Lorenzo-Seva et al., 2011; Ruscio & Roche, 2012). This is why several authors agree on consulting more than one factor-retention criterion (Fabrigar et al., 1999; Goretzko et al., 2021) or using combination rules (e.g., Auerswald & Moshagen, 2019). The pre-trained factor forest model showed very high accuracy in Goretzko and Bühner's (2020) study and may be a more convenient alternative for practitioners as it internally weighs different methods (PA, EKC, CD). As mentioned earlier, the use of model fit indices in factor retention is less defensible (Auerswald & Moshagen, 2019; Montoya & Edwards, 2021). Ultimately, searching for a perfect factor solution in a myriad of available criteria may sometimes be an exercise in futility (e.g., Cattell, 1966). EFA users might, therefore, also consider (in addition to the aforementioned objective recommendations) theoretical utility and a strive towards parsimony in this venture (e.g., Schmitt et al., 2018).

The R packages psych and EFAtools provide functions for the most common criteria (e.g., PA: psych::fa.parallel(efa_data, fm = "wls") or CD: EFAtools::CD (efa_data, n_factors_max = 8)). For our data example, we compared several factor-retention criteria (see https://osf.io/srv8e/ for the full R code). The most reliable methods (see, Auerswald & Moshagen, 2019; Goretzko and Bühner, 2020) – PA, EKC, CD, MAP test, and the factor forest – suggested between five and six factors. Since theoretical considerations (the BFSI claims to measure six facets of the extraversion trait domain with ten items each) speak in favor of a six-factor solution, we retained six latent variables.

Step 5: Factor Rotation

In EFA, all indicators are explained by the set of retained latent factors as dependent variables in a linear regression system. The factor loadings (i.e., the regression

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

parameters) are standardized and expressed in a correlation metric. However, in the initial, unrotated solution, the matrix containing these loadings (factor–indicator correlations) often does not adhere to a clean pattern and makes the interpretation of the factor solution quite difficult. For this reason, factor-rotation methods have been developed to elucidate a more interpretable solution, the so-called *simple structure* (i.e., each indicator loads high on its associated factor and low on all other factors – ideally all cross-loadings are zero) – an idea that was originally presented by Thurstone (1947). More specifically, when estimating the factor loadings and unique variances, the problem of rotation indeterminacy arises (Mulaik, 2010). That is, the loading pattern or loading matrix is only determined up to an arbitrary rotation, and, therefore, selecting an appropriate rotation method solely depends on theoretical considerations and the interpretability of the resulting factor solution. In other words, there is no data-driven way to decide how to rotate the factor solution (see also Browne, 2001; Goretzko et al., 2021).

Orthogonal vs. Oblique Rotation

To obtain such an interpretable solution, two different types of rotations can be used – orthogonal and oblique rotation techniques. Historically, orthogonal rotation methods, which yield uncorrelated factor solutions (all between-factor correlations are constrained to be zero), have been applied more frequently. An advantage of orthogonal or uncorrelated factors is that the respective constructs are clearly distinguishable and that relations between them and third variables can be evaluated independently from each other. However, this process might, in many instances, distort the natural structure of the data when constructs are indeed correlated. Accordingly, oblique rotations, allowing factors to correlate, may be a more appropriate assumption for most social and behavioral research phenomena.

Varimax

The most popular orthogonal rotation method is called varimax (Kaiser, 1958). As the name suggests, the varimax criterion rotates the initial factor solution in a way that maximizes the variance of the squared loadings by columns (i.e., the variance of the squared loadings is maximized for each factor). Hence, this rotation yields rather extreme loadings (either high loadings or very small loadings on each factor).

Quartimax and Equamax

Quartimax is another member of the orthomax family of criteria (Harman, 1976), which includes several orthogonal rotation techniques (e.g., inter alia varimax). Contrary to varimax, it focuses on the row-wise complexity and favors patterns for which each variable has as many zero-loadings as possible. This process leads to an insensitivity to a strong first factor; this is why quartimax often yields a general factor (or a strong first factor and smaller or more trivial second and third factors, etc.). Equamax (see, for example, Kaiser, 1974) is a combination of varimax and quartimax criteria that tries to minimize the number of large loadings per factor and the number of large loadings per variable at the same time.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Promax

One of the most prominent oblique rotation methods is promax (Hendrickson & White, 1964) – a two-stage method that first applies an orthogonal rotation (e.g., varimax) and then subsequently performs the actual oblique rotation. In this process, larger loadings are enhanced compared to smaller loadings by matching the factor loading pattern as closely as possible to an exponentiated version of itself. Usually, the orthogonal loadings are raised to the power of four (e.g., the default setting in the psych package in R; see also the discussion of Hendrickson & White, 1964 on why four was chosen as the default setting for promax), but the exponent can be changed depending on theoretical considerations. It is important to note that a larger exponent will result in larger between-factor correlations.

Oblimin (Family)

Another oblique rotation method that is frequently used in psychological research is called oblimin (Clarkson & Jennrich, 1988). Strictly speaking, it is a family of oblimin methods that includes different rotation techniques, such as quartimin (the oblique generalization of quartimax) or covarimin (the oblique generalization of varimax) as special cases (Clarkson & Jennrich, 1988). A parameter (often named δ [in SPSS] or γ [in R]), which controls the "obliqueness" of the rotated factor solution, determines which rotation of the oblimin family is applied. Jennrich (1979) demonstrated that positive parameter values can be inadmissible when performing oblique rotation; this is why the default value in statistical programs like SPSS and R (we refer to the psych package and the GPArotation package) is zero and corresponds to the quartimin criterion. Quartimin rotation yields a more oblique solution as it minimizes the row-wise complexity by introducing higher inter-factor correlations; decreasing the parameter value (selecting a more negative value) yields a less oblique or more orthogonal solution. Oblimin is selected as the default rotation method in psych.

Geomin

Geomin (Yates, 1987) is a newer oblique rotation technique (there is an orthogonal version as well, e.g., Browne, 2001) that minimizes an objective function based on row-wise geometric means of the squared factor loadings. Thus, geomin focuses on row-wise complexity (i.e., it tries to minimize the number of factors that are needed to explain the variance of each indicator variable). Geomin is the default rotation in Mplus and shows comparably good results when little is known about the true loading pattern (Asparouhov & Muthén, 2009).

Crawford–Ferguson Family

Crawford and Ferguson (1970) presented a general objective function or rotation criterion that is a weighted sum of row and column complexity. Several well-known rotation techniques can be integrated in their general framework. In fact, the

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Crawford–Ferguson (CF) family is equivalent to the orthomax family in the orthogonal case (Crawford & Ferguson, 1970) but yield different results when oblique rotations are considered (Browne, 2001). As a counterpart of quartimax (focus on row-wise complexity), Crawford and Ferguson (1970) introduced CF–facparsim that aims at factor parsimony (focus on column-wise complexity) and, therefore, tries to minimize the number of variables that load on each factor.

Comparison of Rotation Methods

In current research, many EFA users rely on varimax criteron for orthogonal rotation as well as on promax and oblimin/quartimin criteria for oblique rotation (Fabrigar et al., 1999; Goretzko et al., 2021). As pointed out earlier, there is no "correct" way of rotating the initial factor solution. However, many researchers recommend oblique rotation techniques since ruling out between-factor correlations in advance seems to be less plausible in social and behavioral science research (Conway & Huffcutt 2003; Fabrigar et al. 1999; Goretzko et al., 2021).

There are very few recommendations when it comes to choosing an oblique rotation method, though. Simulation studies (e.g., Sass and Schmitt, 2010) suggest that researchers should rely on CF–equamax or CF–facparsim when they expect factor loading patterns with high complexity (i.e., several substantial cross-loadings), whereas geomin or CF–quartimin seem to be more appropriate when patterns closer to simple structure can be assumed. Browne (2001) advocates for trying out different rotation methods (ideally on different subsamples if the data set is large enough for splitting) and to compare the results with regard to stability (if more than one subsample is used) and interpretability. He further suggests comparing a member of the CF family (e.g., CF–equamax) and geomin.

In modern software solutions, a variety of these rotation methods are implemented. The fa function of the psych package, for example, offers numerous options that users can select via the "rotate" argument – psych::fa(efa_data, nfactors = 6, fm = "wls", rotate = "Promax"). For our data example (https://osf.io/srv8e/), we also illustrated the two-step approach – first estimating all parameters for an unrotated solution and then applying a rotation method to increase the interpretability of the factor solution. Comparing the results of orthogonal varimax and oblique quartimin, inter-factor correlations seem to foster interpretability for our exemplary data; in other words, it seems to be reasonable to assume correlated facets, especially as all 60 items are considered to be indicators of the same personality trait (extraversion).

Step 6: Interpretation

The final step in the EFA process is to provide a theoretical interpretation of the solution. Because the analysis is, by definition, exploratory, the theoretical evaluation comes last and is needed to provide meaning to the resulting structure. There are three common vectors of information that EFA users consider in interpretation: factor loadings, communalities, and factor correlations.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

Factor Loadings

The first consideration is the factor loadings; these represent the relationship between the latent factor and an indicator and, more specifically, the degree of variance that the factor accounts for in the indicator. Considering the initial, unrotated factor solution, squared factor loading represents the explained variance in the indicator. The pattern of factor loadings is used to provide meaning to the latent factors in EFA. The indicators with the largest and most distinct loadings are typically considered in the interpretation of the theoretical underpinnings of the latent variable.

There is no universally agreed upon threshold for what constitutes a sufficiently large factor loading for it to be considered meaningful. In psychology, and specifically evaluation of psychological tests at the item level, approximately 0.30–0.40 tends to be considered the lower bound for a meaningful loading; 0.50+ is considered large and substantial (e.g., Gorsuch, 1983). Furthermore, factors that are defined by many cross-loadings (or solely defined by them) are usually not meaningful and signal to the EFA user that too many factors have been extracted or some indicators are poor. A more extreme manifestation of this phenomenon is a "bloated specific" factor (Cattell & Tsujioka, 1964), with one or two very large loadings of indicators on one factor when a broader group of the similar variables are already represented in the remainder of the factor solution. Finally, indicators that have large loadings on more than one factor, unless theoretically indicated as representing variables with clear multiple causes identified in the factor solution (e.g., interstitial variables; see Krueger, 2013, for an example discussion in the personality literature), should also be candidates for elimination as poorly functioning variables (e.g., Brown, 2014).

Communalities

The second consideration on the overall evaluation of a factor solution is the communalities. A communality (h^2) is the total proportion of variance explained in an indicator by all retained factors, whereas $1 - h^2$ is the residual – the proportion of the systematic and unsystematic variance that is unique to the indicator. High communalities typically mean that the factor solution can account for most of the systematic variance in the indicators, whereas low communalities might reflect that the indicators are of lesser importance to the structure being evaluated – these should be considered for removal from the analysis (e.g., Brown, 2014) unless counter-indicated for theoretical reasons (e.g., reduces critical content coverage).

Factor Correlations

As for the final consideration, factor correlations indicate the degree of overlap between the latent factors that have been extracted in the EFA and rotated with an oblique method. These correlations should also be interpreted with theory in mind as there are no thresholds for what constitutes a meaningful correlation. If the emerging latent constructs in an EFA are conceptually expected to be relatively

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

distinct (e.g., positive and negative emotions), smaller correlations are expected. On the other hand, if the constructs are conceptually expected to converge (e.g., impulsivity and risk taking), larger correlations are expected. Extremely high correlations (e.g., 0.80–0.90+) likely reflect significant redundancies in latent constructs and point towards a factor solution with a smaller number of factors.

In our data example, we found that individual items predominantly loaded meaningfully and relatively distinctly onto six latent factors reflecting warmth, gregariousness, assertiveness, drive, adventurousness, and cheerfulness. The latent factors were intercorrelated, as expected, but also distinct (all inter-factor correlations rs < 0.47 when applying quartimin rotation). Communality estimates also indicated that a meaningful proportion of variance was captured in each of the items, with only one exception. Overall, this structure was consistent with theoretical expectations associated with the extraversion trait domain.

Conclusion

This chapter has provided an introduction to the basics of applied EFA. Our goal was to review foundations for and steps associated with conducting an EFA in research. Specifically, we carefully considered each EFA step and described various considerations of which the EFA user should be mindful, including the fact that numerous options and, therefore, researcher degrees of freedom exist for each of these steps. Users should also be aware that many choices they have to make are not only of statistical nature but highly depend on the research questions being addressed, theoretical considerations in general, and the nature of the measured indicators. We believe that EFA is a powerful statistical method for the exploration of higher-order structure, but it is not straightforward and requires many decisions, with the incorrect ones possibly yielding biased results. We hope this guide will therefore be useful to the reader as they choose to apply this method in their research.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–332.

Arendasy, M. (2009) *BFSI: Big-Five Struktur-Inventar (Test & Manual)*. Mödling, Schuhfried GmbH.

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397–438.

Auerswald, M. & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, *24*(4), 468–491. https://doi.org/10.1037/met0000200

Bandalos, D. L. (2008). Is parceling really necessary? A comparison of results from item parceling and categorical variable methodology. *Structural Equation Modeling*, *15*(2), 211–240.

https://doi.org/10.1017/9781009010054.025 Published online by Cambridge University Press

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*(1), 87–101.

Beauducel, A. (2001). On the generalizability of factors: The influence of changing contexts of variables on different methods of factor extraction. *Methods of Psychological Research Online*, *6*(1), 69–96.

Beauducel, A. & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13(2), 186–203.

Ben-Porath, Y. S. & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2 Restructured Form: Manual for Administration, Scoring and Interpretation*. University of Minnesota Press.

Berger, J. L. & Karabenick, S. A. (2016). Construct validity of self-reported metacognitive learning strategies. *Educational Assessment*, *21*(1), 19–33. https://doi.org/10.1080/10627197.2015.1127751

Braeken, J. & van Assen, M. A. (2017). An empirical Kaiser criterion. *Psychological Methods*, 22(3), 450–466. https://doi.org/10.1037/met0000074

Beavers, A. S., Lounsbury, J. W., Richards, J. K., et al. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, *18*(1). https://doi.org/10.7275/qv2q-rk76

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research*. Guilford Press.

Browne, M. W. (1977). Generalized least-squares estimators in the analysis of covariance structures. In D. J. Aigner & A. S. Goldberger (eds.), *Latent Variables in Socio-Economic Models* (pp. 205–226). North-Holland.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150. https://doi.org/10.1207/S15327906MBR3601_05.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*, 476–506.

Cattell, R. B. (1945). The description of personality: Principles and findings in a factor analysis. *The American Journal of Psychology*, *58*(1), 69–90.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cattell, R. B. & Tsujioka, B. (1964). The importance of factor-trueness and validity, versus homogeneity and orthogonality, in test scales. *Educational and Psychological Measurement*, 24(1), 3–30.

Clarkson, D. B. & Jennrich, R. I. (1988). Quartic rotation criteria and algorithms. *Psychometrika*, *53*, 251–259.

Conway, J. M. & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6*(2), 147–168.

Crawford, C. B. & Ferguson, G. A. (1970). A general rotation criterion and its use inorthogonal rotation. *Psychometrika*, *35*, 321–332.

De Winter, J. C. & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*(4), 695–710.

Detrick, P., Ben-Porath, Y. S., & Sellbom, M. (2016). Associations between MMPI-2-RF (restructured form) and Inwald Personality Inventory (IPI) scale scores in a law enforcement pre-employment screening sample. *Journal of Police and Criminal Psychology*, *31*, 81–95.

Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*, *44*(3), 362–388. https://doi.org/10.1080/00273170902938969

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling*, *21*(3), 425–438.

Everitt B. & Hothorn T. (2011) Exploratory factor analysis. In *An Introduction to Applied Multivariate Analysis with R* (pp. 135–161). Springer. https://doi.org/10.1007/978-1-4419-9650-3_5

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272.

Floyd, F. J. & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299.

Goretzko, D. & Bühner, M. (2020). One model to rule them all? Using machine learning algorithms to determine the number of factors in exploratory factor analysis. *Psychological Methods*, *25*(6), 776–786. https://doi.org/10.1037/met0000262

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, *40*(1), 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Gorsuch, R. L. (1983). *Factor Analysis*. Erlbaum.

Graham, J. R., Ben-Porath, Y. S., & McNulty, J. L. (1999). *MMPI-2 Correlates for Outpatient Mental Health*. University of Minnesota Press.

Harman, H. H. (1976). *Modern Factor Analysis*, 3rd ed. University of Chicago Press.

Harman, H. H. & Jones, W. H. (1966). Factor analysis by minimizing residuals (minres). *Psychometrika*, *31*, 351–368. https://doi.org/10.1007/BF02289468

Hendrickson, A. E. & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, *17*(1), 65–70. https://doi.org/10.1111/j.2044-8317.1964.tb00244.x

Henson, R. K. & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*(3), 393–416.

Holzinger, K. J. (1946). A comparison of the principal-axis and centroid factor. *Journal of Educational Psychology*, 37(8), 449–472. https://doi.org/10.1037/h0056539

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve?. *International Journal of Human-Computer Interaction*, *32*(1), 51–62.

Jennrich, R. I. (1979). Admissible values of γ in direct oblimin rotation. *Psychometrika*, *44*, 173–177.

Jokiniemi, K., Pietilä, A. M., & Mikkonen, S. (2021). Construct validity of clinical nurse specialist core competency scale: An exploratory factor analysis. *Journal of Clinical Nursing*, *30*(13–14), 1863–1873.

Jöreskog K. G., Olsson U. H., Wallentin F. Y. (2016) Exploratory factor analysis (EFA). In *Multivariate Analysis with LISREL. Springer Series in Statistics* (pp. 257–282). Springer. https://doi.org/10.1007/978-3-319-33153-9_6

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*, 187–200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). A note on the equamax criterion. *Multivariate Behavioral Research*, *9*(4), 501–503.

Kaiser, H. F. & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, *34*(1), 111–117.

Kirkegaard, E. O. (2016). Some new methods for exploratory factor analysis of socioeconomic data. *Open Quantitative Sociology & Political Science*, *1*(1), November 7. https://doi.org/10.26775/OQSPS.2016.11.07

Krueger, R. F. (2013). Personality disorders are the vanguard of the post-DSM-5.0 era. *Personality Disorders: Theory, Research, and Treatment*, *4*(4), 355–362. https://doi.org/10.1037/per0000028

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949.

Lim, S. & Jahng, S. (2019). Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods*, *24*(4), 452–467. https://doi.org/10.1037/met0000230

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*(3), 285–300. https://doi.org/10.1037/a0033266

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340–364. https://doi.org/10.1080/00273171.2011.564527

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right – Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, *18*(3), 257–284.

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(3), 413–440.

Mulaik, S. A. (2010). *Foundations of Factor Analysis*. CRC Press.

Park, H. S., Dailey, R., & Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. *Human Communication Research*, *28*(4), 562–577.

Pearson, K. (1909). Determination of the coefficient of correlation. *Science*, *30*(757), 23–25.

Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, *48*(1), 28–56. doi:10.1080/00273171.2012.710386

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315.

Ruscio, J. & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, *24*(2), 282–292. https://doi.org/10.1037/a0025697

Sass, D. A. & Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, *45*(1), 73–103. https://doi.org/10.1080/00273170903504810.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*(4), 304–321.

Schmitt, T. A., Sass, D. A., Chappelle, W., & Thompson, W. (2018). Selecting the "best" factor structure and moving measurement validation forward: An illustration. *Journal of Personality Assessment*, *100*(4), 345–362.

Schoedel, R., Au, J. Q., Völkel, S. T., et al. (2018) Digital footprints of sensation seeking. *Zeitschrift Für Psychologie*, *226*(4), 232–245.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Sellbom, M. & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*, *31*(12), 1428–1441. https://doi.org/10.1037/pas0000623

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Thurstone, L. L. (1938). *Primary Mental Abilities*. University of Chicago Press.

Thurstone, L. L. (1947). *Multiple Factor Analysis*. University of Chicago Press.

Thurstone, L. L. & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monographs*, 2, 94.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Wedel, M. & Kamakura, W. A. (2001). Factor analysis with (mixed) observed and latent variables in the exponential family. *Psychometrika*, *66*(4), 515–530.

Widaman, K.F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*, 263–311.

Yates, A. (1987). *Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis*. State University of New York Press.

Yuan, K. H. & Bentler, P. M. (1998). Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*, *51*, 289–309.