Is exploratory factor analysis always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory–exploratory continuum

Article *in* Psychological Methods · May 2023

DOI: 10.1037/met0000579

3 authors:

Pablo Nájera
Universidad Pontificia Comillas

Francisco J Abad
Universidad Autónoma de Madrid

Miguel A. Sorrel
Universidad Autónoma de Madrid

Is Exploratory Factor Analysis Always to be Preferred? A Systematic Comparison of Factor Analytic Techniques throughout the Confirmatory-Exploratory Continuum

Pablo Nájera, Francisco J. Abad, and Miguel A. Sorrel
Department of Social Psychology and Methodology, Universidad Autónoma de Madrid

The official citation that should be used for this material is:

Nájera, P., Abad, F. J., & Sorrel, M. A. (2023). Is exploratory factor analysis always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory–exploratory continuum. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000579

Author Note

Pablo Nájera https://orcid.org/0000-0001-7435-2744

Francisco J. Abad https://orcid.org/0000-0001-6728-2709

Miguel A. Sorrel https://orcid.org/0000-0002-5234-5217

This research was partially supported by Ministerio de Ciencia, Innovación y Universidades of Spain (Grant PSI2017-85022-P), European Social Fund, and Cátedra de Modelos y Aplicaciones Psicométricas (Instituto de Ingeniería del Conocimiento and Universidad Autónoma de Madrid). Portions of these findings were presented at the 9th European Congress of Methodology (July 2021; Valencia, Spain) and X Jornadas AIIDI (December 2021; Madrid, Spain). The data of this research has been made publicly available at https://psychometricmodelling.shinyapps.io/FAcomparison/, and the simulation codes and methods used can be accessed at https://osf.io/vw4xn/. This study was not preregistered. We have no conflict of interest to disclose.

Corresponding concerning this article should be addressed to Pablo Nájera, Faculty of Psychology, 6 Iván Pavlov St, Cantoblanco Campus, Madrid, Spain, 28049. Email: pablo.najera@uam.es

Abstract

The number of available factor analytic techniques has been increasing in the last decades. However, the lack of clear guidelines and exhaustive comparison studies between the techniques might hinder that these valuable methodological advances make their way to applied research. The present paper evaluates the performance of confirmatory factor analysis (CFA), CFA with sequential model modification using modification indices and the Saris procedure, exploratory factor analysis (EFA) with different rotation procedures (Geomin, target, and objectively refined target matrix), Bayesian structural equation modeling (BSEM), and a new set of procedures that, after fitting an unrestrictive model (i.e., EFA, BSEM), identify and retain only the relevant loadings to provide a parsimonious CFA solution (ECFA, BCFA). By means of an exhaustive Monte Carlo simulation study and a real data illustration, it is shown that CFA and BSEM are overly stiff and, consequently, do not appropriately recover the structure of slightly misspecified models. EFA usually provides the most accurate parameter estimates, although the rotation procedure choice is of major importance, especially depending on whether the latent factors are correlated or not. Finally, ECFA might be a sound option whenever an a priori structure cannot be hypothesized and the latent factors are correlated. Moreover, it is shown that the pattern of the results of a factor analytic technique can be somehow predicted based on its positioning in the confirmatory-exploratory continuum. Applied recommendations are given for the selection of the most appropriate technique under different representative scenarios by means of a detailed flowchart.

Keywords: confirmatory factor analysis, exploratory factor analysis, Bayesian structural equation modeling, cross-loadings, internal structure.

Is Exploratory Factor Analysis Always to be Preferred? A Systematic Comparison of Factor Analytic Techniques throughout the Confirmatory-Exploratory Continuum

Factor analysis is arguably the most used statistical tool to examine the internal structure of scales and questionnaires in psychological and educational assessment. Not surprisingly, factor analysis is still a topical issue after more than a century since its original formulation. This is especially noticeable in the number of factor analytic techniques that have been proposed in the last decades. Apart from the traditional exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), applied researchers can choose from a wide range of recently available methods and variations such as different sequential model modification procedures (e.g., Marcoulides et al., 1998; Saris et al., 2009), exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) with different rotation procedures (Browne, 2001) including, for example, the objectively refined matrix procedure (RETAM; Lorenzo-Seva & Ferrando, 2020), Bayesian structural equation modeling (BSEM; Muthén & Asparouhov, 2012), Set-ESEM (Marsh et al., 2020), regularized structural equation modeling (Jacobucci et al., 2016), and penalized likelihood structural equation modeling (Huang et al., 2017), among others.

With all these options on the table, it might be difficult to decide which method is most appropriate for a given dataset. To shed some light on the question, a few simulation studies have been conducted to compare the performance of some factor analytic techniques (Asparouhov & Muthén, 2009; Guo et al., 2019; Liang et al., 2020; Muthén & Asparouhov, 2012; Xiao et al., 2019; Whittaker, 2012). In general, these studies concluded that: (a) CFA is usually overly restrictive and leads to poor model fit and biased parameter estimates (particularly an overestimation of factor correlations), (b) EFA/ESEM often obtains satisfactory model fit and provides more accurate parameter estimates, (c) BSEM obtains accurate parameter estimates as long as the priors are correctly specified, and (d) sequential model modifications usually perform poorly at recovering the generating internal structure. Based on these findings, it might seem that the selection of the most appropriate technique is an already answered question: EFA/ESEM would be the indisputable winner, perhaps followed by BSEM only if the priors are correctly specified. This notion has made an impression on applied researchers, and several scales that were already validated by means of CFA have been reanalyzed using EFA/ESEM or BSEM, leading to solutions with better model fit and substantially lower factor correlations (e.g., Arens & Morin, 2016; Fong & Ho, 2013; Garrido et al., 2020).

However, we can identify some questions that make the selection of the most appropriate factor analysis technique still an open issue. First, all the aforementioned simulation studies compared only a small number of techniques, such as sequential model modifications in CFA (Whittaker, 2012), CFA and ESEM (Asparouhov & Muthén, 2009), CFA and BSEM (Muthén & Asparouhov, 2012), ESEM and BSEM (Liang et al., 2020), or CFA, ESEM, and BSEM (Guo et al., 2019; Xiao et al., 2019). Second, a limited set of conditions was considered, including a fixed sample size (Xiao et al., 2019), number of factors (Asparouhov & Muthén, 2009; Guo et al., 2019; Liang et al., 2020; Muthén & Asparouhov, 2012; Whittaker, 2012), factor correlations (Guo et al., 2019; Liang et al., 2020; Muthén & Asparouhov, 2012), or major loadings magnitude (Asparouhov & Muthén, 2009; Guo et al., 2019; Liang et al., 2020; Muthén & Asparouhov, 2012; Xiao et al., 2019). The number of items per factors was also fixed in all these studies. Third, the performance of the techniques was evaluated almost exclusively based on model fit and parameter estimation accuracy. However, the determinacy of factor score estimates has profound implications in the definition of the latent constructs, that is, validity (Grice, 2001). In this vein, Booth and Hughes (2014) observed that, despite the differences in model fit and parameter estimates between CFA and EFA, both techniques showed very similar factor score estimates.

Given all this, the present paper aims to fill these gaps by comparing the performance of several factor analytic techniques under a unified set of conditions by means of a Monte Carlo simulation study. Apart from several already existing techniques, a new procedure that combines EFA (or BSEM) and CFA is here presented and evaluated. The ultimate goal of the simulation study is to provide applied recommendations and guidelines to help researchers choose the most appropriate technique as a function of the model and data characteristics.

The remainder of the paper is laid out as follows. First, a brief introduction of EFA, CFA, and some middle-ground solutions between these approaches is presented. Second, a procedure that systematically combines EFA (or BSEM) and CFA is introduced. Third, the performance of the factor analytic techniques is evaluated by means of a simulation study. Fourth, the techniques are illustrated using real data. Finally, a discussion is made including applied guidelines and implications, as well as future research lines.

Exploratory Factor Analysis

EFA was developed in the first half of the 20th century, based on the work of Spearman and Thurstone, thus being the first available factor analytic technique. This was for a time the only alternative to explore the internal structure of scales and questionnaires. With the emergence of the CFA in the 1970s (described below), which provided some theoretical and technical advantages, EFA became more oriented towards the evaluation of new instruments, while CFA gained popularity as a tool to examine the structure of instruments that had already received support in the previous literature. Nonetheless, the recent development of ESEM, which allows for the specification of EFA within structural equation modeling, has rekindled the interest of both applied and methodological studies in EFA.

Arguably, the defining feature of EFA is that, as an unrestricted, data-driven procedure, all loadings can obtain non-zero values. Thus, by estimating all possible cross-loadings, this technique is not prone to model misspecifications (beyond dimensionality issues). This usually results in good model fit and accurate parameter estimates (Asparouhov & Muthén, 2009; Marsh et al., 2014; Perry et al., 2015; Sorrel et al., 2021). These desirable features, however, come at some costs. First, EFA suffers from rotational indeterminacy, that is, there exist infinite factor loading matrices in EFA that are compatible with the observed covariance matrix. To minimize the degree of arbitrariness, several rotation procedures have been developed aiming to find a loading structure that facilitates interpretation among the infinite possible options. Even though each rotation procedure minimizes a particular function, often leading to substantially different rotated parameters (see Browne, 2001; Sass & Schmitt, 2010; Schmitt & Sass, 2011), they usually seek a simple structure. Rotation procedures are usually divided into orthogonal procedures (e.g., Varimax; Kaiser, 1958), which force factors to be independent, and oblique procedures (e.g., Geomin; Yates, 1987; Promin; Lorenzo-Seva, 1999; Oblimin; Clarkson & Jennrich, 1988), which allow factors to be inter-correlated. Oblique rotation is usually preferred in psychological sciences due to the inherently complex and interrelated nature of the latent constructs (Goretzko et al., 2021). Furthermore, the great complexity of EFA models (i.e., the estimation of many parameters) usually demand large sample sizes, as EFA shows estimation issues and unstable results with small samples ($N < 300$; Liang et al., 2020; Marsh et al., 2020). This problem is not exclusive for one estimation method (e.g., Liang et al., 2020; Ximenez, 2006), although maximum likelihood can be more prone to it in the presence of weak factors and very small sample sizes ($N = 100$; Briggs & MacCallum, 2003). Moreover, as any other data-driven method would, EFA is also prone to capitalization on chance because the final rotated loading matrix will inherently contain some degree of sampling error that does not reflect the actual population model. The extent to which the amount of error might cause a disruption to the substantial interpretation of the model will decrease with larger sample sizes, but the final EFA solution should be always evaluated in terms of theoretical plausibility. In this vein, just as a meaningless parameter should not be included in a CFA just to improve model fit (as discussed in the next section), an untenable cross-loading should not be retained in an EFA. The blind acceptance of EFA solutions that are difficult to explain will inevitably lead to models that are difficult to generalize. This issue is of special relevance from an applied perspective because the satisfactory model fit often showed by EFA might tempt practitioners to retain these models without further questioning.

The above description refers to EFA with a mechanical rotation procedure. Mechanical rotation procedures (e.g., Varimax, Geomin) are purely data-driven procedures and, thus, they do not require the researcher to prespecify the internal structure beyond determining the number of factors. Contrary to these methods, target rotation incorporates theoretical knowledge into EFA by means of a prespecified target matrix that reflects the hypothesized internal structure. By allowing factors to correlate (oblique target rotation) or not (orthogonal target rotation), this procedure will provide, among the infinite possible loading matrices, the one that is closer to the target matrix. Thus, it is regarded as a middle-ground solution between EFA and CFA (Browne, 2001; Marsh et al., 2014). Due to this balance, EFA with target rotation has been considered as a sound option to enhance interpretability and reduce the probability of capitalization on chance (Marsh et al., 2014). However, it should be noted that, unlike mechanical procedures, target rotation is subject to misspecification errors (Lorenzo-Seva & Ferrando, 2020). Although the consequences of such errors are expected to be milder than those in a CFA model, model misspecification might result in biased parameters.

Confirmatory Factor Analysis

CFA was developed in the 1970s as a restricted version of EFA with the purpose of conducting factor analysis from a hypothesis-testing approach. Accordingly, CFA requires an *a priori* specification of the internal structure of the questionnaire, which is then tested by fitting the model to the data. In practice, an independent clusters model CFA (ICM-CFA) is often specified by freely estimating only one target loading (i.e., theoretically expected loading) per item, while fixing to zero the remaining non-target loadings. This makes ICM-CFA a purely hypothesis-driven approach that focuses on the substantial meaning and interpretation of the model. However, due to the inherent complexity of most psychological constructs, it has been argued that it is virtually impossible to create pure indicators that load on a single factor (Asparouhov & Muthén, 2009; Booth & Hughes, 2014; Hopwood & Donnellan, 2010; Marsh et al., 2014). Instead, small but non-negligible cross-loadings are expected at the population level, which will cause CFA to fail in terms of poor model-data fit and biased parameter estimation, particularly with an overestimation of factor correlations (Asparouhov & Muthén, 2009; Xiao et al., 2019).

To mitigate these problems, researchers often make sequential model modifications to the CFA by freeing parameters that are likely to be misspecified. The inspection of modification indices (MI; Satorra, 1989; Sörbom, 1989) and expected parameter change (EPC; Saris et al., 1987) serve for this matter. The MI associated to a fixed parameter indicates the gain in model fit if the parameter were freely estimated, while the EPC is a direct estimation of the expected value of such parameter if freed. Hence, an ICM-CFA can become more flexible by introducing the cross-loadings with a large associated MI or EPC in the model. As a counterpart, the better model fit can come at the cost of capitalization on chance. That is, these data-driven model modifications can be based on idiosyncratic characteristics of a particular sample, thus being non-generalizable to subsequent samples (Browne, 2001; MacCallum et al., 1992). This issue can be further aggravated if model misspecifications are present in the original formulation of the ICM-CFA. The recommended strategy to mitigate this problem is to make model modifications only if the sample size is sufficiently large to provide stable results (usually over 1,000), the number of suggested modifications is not large and, most importantly, such modifications have a meaningful interpretation according to theory. The relevance of these precautions is even higher as simulation studies have reported a non-optimal performance of both MI and EPC in consistently identifying misspecified parameters (Whittaker, 2012; Yuan & Liu, 2021). Lastly, even though the present paper focuses on scenarios with a single sample, it should be emphasized that using two or more samples to cross-validate the structure found with one particular sample using either EFA or CFA is always recommended as a complement to the evaluation of the theoretical interpretability, especially to check whether unexpected findings are generalizable or can be regarded as spurious (MacCallum et al., 1992).

Between Confirmatory and Exploratory: From Theory to Data

As stated above, the implementation of either a purely confirmatory or purely exploratory model comes with a host of methodological and substantive challenges. To mitigate them, several techniques have been recently developed with the aim of complementing theoretical and empirical information. These new proposals can be regarded as middle-ground solutions between the confirmatory and exploratory poles of the factor analysis continuum.[1] Two traditional middle-ground solutions have been already mentioned: (a) the use of MI or EPC to make sequential modifications in CFA models, and (b) the implementation of EFA with target rotation. There are, however, two methods related to these solutions that are worth mentioning. First, Saris et al. (2009) developed a more comprehensive sequential model modification procedure by considering the MI, the statistical power associated to the MI, and the fully standardized EPC (Chou & Bentler, 1993) to decide whether a parameter should be freed. Although the Saris procedure relies on the somehow

[1] It has been repeatedly stated in the literature that the confirmatory and exploratory terms are not precise in the sense that a CFA can be implemented in an exploratory fashion, while an EFA can be implemented in a more confirmatory fashion (e.g., Browne, 2001). We agree that the terms *restricted* and *unrestricted* factor analysis (Ferrando, 2021) are far more accurate, and the only reason why we stick to the conventional terminology is to facilitate readability.

arbitrary selection of cutoff values for each index, it is an exhaustive method that can potentially mitigate the limitations associated to the sample size dependency of MI. However, despite this promising feature, Whittaker (2012) found that the Saris procedure did not consistently provide a more accurate identification of relevant misspecified parameters than sole consideration of MI. Second, Lorenzo-Seva and Ferrando (2020) recently developed a method to apply sequential modifications to a target matrix, so that the final EFA solution is based on an objectively refined target matrix (RETAM). The RETAM procedure is closely related to the iteration target rotation algorithm previously developed by Moore et al. (2015). The main difference between both methods, and the reason why we focus here on RETAM, is that the iteration target rotation algorithm is a purely exploratory procedure that is initiated by using a standard analytic rotation method (which will be later modified in an iterative manner), while RETAM requires researchers to specify a target matrix in advance, thus providing the method a confirmatory basis (Lorenzo-Seva & Ferrando, 2020). Specifically, in Step 1 of the RETAM procedure, an EFA using a partially specified target rotation is applied. Note that the target matrix includes some zeros, which indicate that the researcher anticipates that the item will not have a salient loading on the factor. The remaining entries of the target matrix are unspecified (i.e., with no value), and thus the target rotation only minimizes the sum of the squared rotated loadings corresponding to the zero positions. In step 2, the Promin approach is applied to the target rotated matrix to refine it. Specifically, the target rotated matrix is row-normalized, and a cutoff is computed for each column (i.e., $m_k + s_k/4$, where m_k and s_k are the mean and standard deviation of the squared loadings in column *k*). The loadings whose squared magnitude is lower than their corresponding cutoff will be the zeros in the new target matrix. Steps 1 and 2 are repeated until the partially specified target matrix does not change in two consecutive iterations. The authors call it *objectively refined target matrix* since the refinement process is based solely on empirical information (without the involvement of human judgment). Lorenzo-Seva and Ferrando (2020) showed that the RETAM procedure was able to correct misspecification errors in the target matrix and, thus, provide more accurate parameter estimates. The performance of RETAM under different conditions than those covered in their study (e.g., 20 items and 3 independent factors) is yet to be evaluated.

Another recent technique that has gained considerable attention is BSEM. BSEM is a middle-ground solution because the internal structure of the questionnaire needs to be specified a priori (like in CFA), but no loading is fixed to zero (like in EFA). So far, BSEM resembles EFA with target rotation. The difference between both methods is that in BSEM the hypothesized model formulation is incorporated in the estimation by using prior distributions for the factor loadings. Namely, informative priors (e.g., normal distributions with small variance) are used for non-target loadings, while non-informative priors (e.g., uniform distributions or normal distributions with large variance) are used for major loadings (Muthén & Asparouhov, 2012). Thus, BSEM gives researchers more flexibility when incorporating prior knowledge into the model formulation. Motivated by these desirable characteristics, BSEM has been already used in scale validation studies (e.g., Fong & Ho, 2013; Golay et al., 2013). Moreover, it has shown promising results in recent simulation studies with continuous variables (Guo et al., 2019; Xiao et al., 2019; Wei et al., 2022) and categorical variables (Liang et al., 2020), as well as bifactor models (Zhang et al., 2021). Overall, these studies point out that BSEM might be a suitable option under small sample size conditions, particularly when cross-loadings are not large. However, simulation results also highlight that the specification of the priors has a great impact on BSEM results. For instance, Xiao et al. (2019) found that BSEM with a correct prior specification clearly outperformed EFA in parameter estimation accuracy, but this was not the case for BSEM with priors with mean zero for cross-loadings. Guo et al. (2019) also obtained substantial differences in

BSEM solutions depending on whether a prior with mean 0 or 0.1 was used for cross-loadings. In practice, correctly specifying the priors is virtually impossible (MacCallum et al., 2012; Xiao et al., 2019), and the implementation of sensitivity analyses to evaluate the consistency of BSEM results under different prior specifications has been repeatedly recommended (e.g., Liang et al., 2020; MacCallum et al., 2012; Muthén & Asparouhov, 2012; Xiao et al., 2019). Implementing such a sensitivity analysis might be a difficult task in applied studies. Varying the variance of the priors seems easier than applying different means to different loadings; however, it has been shown that the mean of the priors has a much greater effect on the results compared to the variance of the priors (Liang et al., 2020; Xiao et al., 2019; Wei et al., 2022). An additional difficulty for practitioners might be the evaluation of model fit in BSEM analyses; until recently, the traditional goodness-of-fit indices that are available for CFA and EFA could not be computed for Bayesian analyses (MacCallum et al., 2012; Rindskopf, 2012). Finally, some authors have pointed out that, similarly to EFA, the estimation of all cross-loadings in BSEM can lead to complex structures that, subsequently, result in implausible and non-generalizable models (Rindskopf, 2012; Stromeyer et al., 2015).

Between Confirmatory and Exploratory: From Data to Theory

All the procedures described above are based on a previously hypothesized simple internal structure, but then allowed to achieve more complex patterns based on the empirical data. That might be considered as a desirable feature, since it lets researchers testing their specific theories, acknowledging the possibility of some deviations (i.e., cross-loadings). On the other hand, relying on a hypothesized internal structure is always inherently associated to a certain probability and magnitude of model misspecification (Yuan & Liu, 2021), which can be a cause of concern. For instance, a moderate amount of misspecification errors in the original model will increase the probability of making an early mistake in sequential modifications, disrupting subsequent ones (Yuan & Liu, 2021). In this case, the complex final structure would be uninterpretable and unstable (MacCallum et al., 1992; Marsh et al., 2020).

Given the above, a “free-to-restricted” middle-ground solution that can achieve a parsimony solution while being free of potential model misspecification is still to be developed. Here, we propose a method that can fulfil these criteria. The procedure is conceptually simple and, despite having been tangentially discussed in the literature (e.g., Bandalos & Finney, 2019; Schmitt et al., 2018), to the authors’ knowledge it has not been implemented in applied studies nor systematically evaluated in simulation studies. The broad rationale of the procedure consists in three steps: (a) fit an EFA model to the data using a mechanical rotation procedure, (b) identify the relevant loadings from the rotated EFA solution, and (c) fit a CFA model by fixing to zero the non-relevant loadings and freely estimating the relevant loadings. This method is referred to as EFA-based CFA (ECFA). Although the idea of introducing restrictions to unrestricted factor analysis is not new (e.g., Huang et al., 2017), we propose a new set of procedures to implement it in a systematic fashion.

The first step of the ECFA is equivalent to a conventional EFA implementation and, accordingly, researchers need to specify the extraction method and the rotation procedure (see Goretzko et al., 2021, and Izquierdo et al., 2014, for accessible reviews about EFA uses and recommendations). In the remainder of this paper, maximum likelihood estimation and oblique Geomin rotation as implemented in *Mplus* version 8 (Muthén & Muthén, 2017) will be assumed for the first step of the ECFA procedure. Please note that, in the codes provided below, the ECFA method can be also implemented using the `lavaan` package (Rosseel, 2012) from R software (R Core Team, 2021).

Secondly, in order to develop an empirical procedure to identify the relevant loadings, an operative definition of *relevant* loadings in first required. Two definitions are here considered. The first one defines a relevant loading as a loading that is different from zero in the population, that is, a statistically significant loading. The use of factor loading standard errors to get information about the stability of the internal structure is not new, but has been widely overlooked (Asparouhov & Muthén, 2009; Cudeck & O'Dell, 1994; Schmitt & Sass, 2011). The third step of the ECFA procedure would then consist in fitting a CFA in which the statistically significant loadings from the EFA solution are freely estimated, while the non-statistically significant loadings are fixed to 0. This variation, based on the *p*-value, will be referred to as $ECFA_p$.

Despite the advantages of using standard errors to identify the relevant factor loadings, the dependency on sample size might be an issue under some scenarios, such as low-scale assessments. It this context, small cross-loadings might be non-significant despite having a substantial impact on parameter estimation (Asparouhov & Muthén, 2009). This issue is addressed by the second definition of relevant loading, which determines the importance of the loadings based on their relative contribution to the variance of the item. We propose the R-squared method, which is aligned with this definition. The R-squared method is based on the work of de la Torre and Chiu (2016) on the empirical Q-matrix validation in the cognitive diagnosis modeling framework. To understand the rationale of the method, let J denote the number of items and K denote the number of factors. Note that there is a total of 2^K possible configurations when determining what factors are being measured by an item. Following the terminology used in cognitive diagnosis modeling, we will denote the binary vector of factors being measured by an item as a *q-vector*. For instance, the q-vector of an item measuring the second and third factors in a four-dimensional test (K = 4) would be q = {0, 1, 1, 0}. Furthermore, let $K^{(m)}$ denote the number of factors included in a q-vector and $\boldsymbol{k}^{(m)}$ denote the vector that identifies the position of such factors ($K^{(m)} = 2$ and $\boldsymbol{k}^{(m)} = \{2, 3\}$ in the previous example). The pseudo-algorithm for the complete implementation of the $ECFA_{R2}$

(i.e., the ECFA using the R-squared method for relevant loading identification) procedure is presented in the following:

1. Fit an EFA with a mechanical oblique rotation procedure (e.g., Geomin).
2. Extract the structure matrix **S**:

$$\mathbf{S} = \mathbf{\Lambda\Phi}, \tag{1}$$

   where $\mathbf{\Lambda}$ is the (rotated) factor loading matrix and $\mathbf{\Phi}$ is the factor intercorrelation matrix. Each entry in **S** (s_{jk}) corresponds to the correlation between item *j* and factor *k*.
3. Calculate the proportion of variance accounted for each item *j* by the factors included in each q-vector *m* (R^2_{jm}):

$$R^2_{jm} = \boldsymbol{s}_{j\boldsymbol{k}^{(m)}} \mathbf{\Phi}^{-1}_{\boldsymbol{k}^{(m)}\boldsymbol{k}^{(m)}} \boldsymbol{s}^{\mathrm{T}}_{j\boldsymbol{k}^{(m)}}. \tag{2}$$

   Note that $R^2_{jm} = \boldsymbol{s}^2_{j\boldsymbol{k}^{(m)}}$ for the q-vectors with $K^{(m)} = 1$, and $R^2_{jm} = h^2_j$ for the q-vector with $K^{(m)} = K$, where h^2_j is the communality of item *j*.
4. For each item *j* and q-vector *m*, define the *proportion of variance accounted for* (PVAF_{jm}) as

$$\text{PVAF}_{jm} = \frac{R^2_{jm}}{h^2_j}. \tag{3}$$

5. For each item *j*, define a *candidate q-vector* for each $K^{(m)} \in \{1, \ldots, K\}$ as the one with the highest R^2_{jm} among the q-vectors with the same $K^{(m)}$. There will be *K* candidate q-vectors per item, one for each $K^{(m)}$.
6. For each item *j*, define the set of *appropriate q-vectors* as those candidate q-vectors that fulfill $\text{PVAF}_{jm} > \varphi$, where φ is a prespecified cutoff point.
7. For each item *j*, define the *suggested q-vector* as the simplest q-vector (i.e., the one with the lowest $K^{(m)}$) among the set of appropriate q-vectors.
8. Fit a CFA by freely estimating the relevant factor loadings (identified as 1 in the suggested q-vectors) and fixing to zero the non-relevant factor loadings (identified as 0 in the suggested q-vectors).

In summary, the R-squared method consists in: (a) fitting an EFA to the data, (b) identifying which factors are relevant for each item (i.e., which combination of factors explain a large proportion of the item's variance), and (c) fitting a CFA to the data based on the identified relevant loadings. Differently from the use of traditional cutoff values directly applied to each individual loading (e.g., $\lambda \geq 0.3$), the R-squared method applies a more comprehensive approach by considering all the loadings of an item, in addition to factor correlations, to determine its most appropriate q-vector. Table 1 illustrates the functioning of the $ECFA_{R2}$, showing the suggested q-vector for item *j* as a function of the factor correlation (ϕ_{12}), the cross-loading magnitude (λ_{j1}), and the major loading magnitude (λ_{j2}). It should be noted that the choice of φ will have an important role in the final suggested q-vectors. Higher values for φ will result in more complex models. In the extreme case of $\varphi = 1$, all suggested q-vectors will be fully specified (i.e., $\{\mathbf{1}\}$), resulting in a completely unrestricted model. On the opposite, a low value for φ will result in simpler models, to the point where a very low φ will lead to an ICM-CFA formulation. Researchers can specify the value of φ either based on their desired minimum proportion of variance accounted for the items by the factors or an empirical criterion. Namely, the latter approach can be accomplished by following the one used in penalized likelihood SEM (Huang et al., 2017; Huang, 2020), where different values of φ can be employed to fit several CFA based on the different resulting model formulations, and then the optimal model can be selected as a function of a model fit index. The Bayesian information criterion (BIC; Schwarz, 1978) is here selected following Huang et al. (2017).

[Please insert Table 1 here]

Finally, note that either the *p*-value criterion or the R-squared method can be applied to other factor analytic techniques apart from EFA. For instance, if a researcher wanted to incorporate previous knowledge into the initial internal structure while achieving a parsimonious final solution, these approaches could be easily incorporated to BSEM, resulting in the BSEM-based CFA (BCFA) procedure. In addition to $ECFA_p$ and $ECFA_{R2}$, we also explore the performance of $BSEM_p$ and $BSEM_{R2}$ in the present study.

The Present Investigation

Addressing the limitations of previous research focused on the comparison of several factor analytic techniques, the present study provides a broad comparison framework covering a larger set of methods, simulation conditions, and performance measures. Namely, ICM-CFA, CFA with sequential model modifications using either MI (CFA_{MI}) or the Saris procedure (CFA_S), EFA with Geomin (EFA_G) and target (EFA_T) rotation, the RETAM procedure, BSEM, $ECFA_p$, $ECFA_{R2}$, $BCFA_p$, and $BCFA_{R2}$ will be systematically evaluated. Furthermore, contrasting with previous simulation studies (with the only exception of Lorenzo-Seva & Ferrando, 2020), in the present research the data were generated in a more realistic way by introducing misfit in the population covariance matrix (Cudeck & Browne, 1992; Lai, 2019). The ultimate goal of the simulation study, which is complemented with a real data illustration, is to serve as the basis for the development of applied recommendations and guidelines.

To facilitate the formulation of specific research questions and hypotheses, Table 2 summarizes the main characteristics of the eleven factor analytic techniques to be compared in the present study. The techniques have been ordered based on a confirmatory-exploratory continuum as found in the simulation study (this is described in the Similarity of the Factor Analytic Techniques subsection). First, regarding the initial specification of the internal structure, hypothesis-driven techniques will facilitate the substantial interpretation of the final solution, although at the potential cost of having model misspecifications. On the contrary, data-driven procedures are aimed at finding an approximately simple structure by optimizing a complexity function of factor loadings, which might increase the risk of lacking theoretical interpretability. Second, factor analytic techniques with a restricted factor space will produce more parsimonious loading matrices, which might facilitate the understanding of the final solution; however, if non-negligible cross-loadings have been incorrectly fixed to 0, the model will lack flexibility to accommodate such misspecification, leading to a worse model fit and biased parameter estimates. On the other hand, unrestricted techniques will be able to address this problem by giving non-zero values to small factor loadings; as a counterpart, the resulting complex solutions might hinder the interpretability of the model. Third, the *original* techniques (i.e., ICM-CFA, BSEM, EFA_T, EFA_G) can be distinguished from their *variants* (i.e., CFA_{MI}, CFA_S, $BCFA_p$, $BCFA_{R2}$, RETAM, $ECFA_p$, $ECFA_{R2}$), which apply modifications to the structure obtained by their original technique. Following the categorization of Lorenzo-Seva and Ferrando (2020), model modifications can be made by only considering the inclusion of new parameters (make more complex), only considering the removal of already-included parameters (make simpler), or both (complete refinement). According to this categorization, the procedures that aim to make restrictive models more complex (e.g., CFA_{MI}, CFA_S) are expected to be more dependent to the initial degree of misspecification than the procedures that aim to simplify unrestricted models (e.g., ECFA). The specific research questions and hypotheses to be explored in the present study are enumerated in the following:

(Q1) *How can the factor analytic techniques be grouped or categorized according to their functioning across conditions when the model (i.e., initial internal structure) is slightly misspecified?* A broader understanding of the similarity between the performance of the methods is expected to be achieved by comparing all the above-mentioned techniques under a unified set of conditions. Namely, ICM-CFA is expected to perform very differently (and objectively worse) compared to the other procedures. EFA_T is expected to be located between EFA_G and BSEM in terms of functioning. The variants of the original procedures (i.e., CFA_{MI}, CFA_S, $BCFA_p$,

$BCFA_{R2}$, RETAM, $ECFA_{p}$, $ECFA_{R2}$) are expected to behave more similarly between each other compared to the original procedures (i.e., ICM-CFA, BSEM, EFA_{T}, EFA_{G}).

(Q2) *When the baseline model is theoretically constrained (i.e., ICM-CFA) but slightly misspecified, does sequential model modification substantially improve the recovery of the internal structure of a test?* Previous studies have reached unclear conclusions regarding the use of MI, stating that they can be provide useful information but should not be solely used for model modification (Whittaker, 2012). Considering a slightly misspecified model, we expect CFA_{MI} and CFA_{S} to provide a better recovery of the internal structure compared to ICM-CFA.

(Q3) *When the baseline model is unconstrained but theoretically guided (e.g., BSEM, EFA_{T}) and slightly misspecified, does adding model constraints (i.e., BCFA) or refining the initial internal structure (i.e., RETAM) provide substantial advantages?* The RETAM procedure has been only evaluated in the original study by Lorenzo-Seva and Ferrando (2020), while the BCFA techniques have not been previously studied. BCFA procedures are expected to increase model stability by fixing to zero the non-relevant factor loadings resulting from a BSEM, while retaining the relevant cross-loadings. RETAM is expected to identify the non-zero cross-loadings, thus providing more accurate parameter estimates compared to EFA_{T}.

(Q4) *When the baseline model is unconstrained and not guided by the theory (i.e., EFA_{G}), does adding model constraints (i.e., ECFA) provide substantial advantages?* ECFA procedures are expected to differentiate the negligible cross-loadings from the non-zero cross-loadings based on the EFA_{G} solutions, leading to a more accurate parameter estimation.

(Q5) *Do the hypothesis-based techniques (i.e., by fixing, targeting, or setting priors for expected zeros) substantially improve the performance of the theoretically blind techniques (i.e., EFA_G, ECFA) when the model is only slightly misspecified?* The overly restrictive structure of ICM-CFA is expected to be unable to accommodate slightly misspecified models. The remaining hypothesis-driven techniques are expected to perform similarly to the data-driven techniques under slightly misspecified models, due to either their ability to identify relevant cross-loadings or to be flexible enough to accommodate them.

(Q6) *Is there a clear correspondence between parameter estimation accuracy, model fit indicators, and determinacy of factor score estimates?* This correspondence is not expected. On the one hand, sequential model modifications are widely known for improving model fit despite not always making the proper adjustments (MacCallum et al., 1992). On the other hand, it has been shown that EFA and CFA can provide similar factor score estimates despite showing very different factor loadings and model fit (Booth & Hughes, 2014).

[Please insert Table 2 here]

Simulation Study

Design

A Monte-Carlo simulation study was conducted to evaluate the performance of eleven factor analytic techniques by manipulating seven simulation factors: sample size (N = 300, 650, 1000), number of factors (K = 3, 5), number of indicators per factor[2] (JK = 4, 8), factor correlations (FC = 0, 0.5), magnitude of main loadings (ML = 0.5, 0.7), magnitude of cross-loadings (CL = 0.15, 0.30), and number of cross-loadings per factor (CLK = 1, 2). The levels

[2] By *number of indicators per factor*, we refer to the number of items that have a major loading (i.e., with a magnitude equal to *ML*) on each factor.

for each simulation factor were selected in pursue of representativeness of applied settings, reflecting a small/weak or large/strong condition. For instance, even though many scale validation studies make use of large sample sizes ($N > 1{,}000$; e.g., Arens & Morin, 2016; Wiesner & Schanding, 2013), more accessible smaller samples ($N < 350$) are also common (e.g., Fong & Ho, 2013; Golay et al., 2013). In addition, many popular psychological scales such as the NEO-FFI (Costa & McCrae, 1992), the Big Five Inventory (BFI; John et al., 1991), or the Personality Inventory for DSM-5 (PID-5; Krueger et al., 2012) are designed to measure around 3 and 5 factors with a number of items/facets per factor that usually varies between 4 and 10 (e.g., Fong et al., 2015; Garrido et al., 2020). Moreover, substantive loadings are often found to fall within a range from 0.5 to 0.8 (e.g., DiStefano et al., 2017; Perry et al., 2013), while moderate to strong factor correlations up to 0.5 are commonly encountered in personality or clinical scales (e.g., Booth & Hughes, 2014; Wiesner & Schanding, 2013). Finally, even though the amount and magnitude of cross-loadings greatly depends on the context of study, it is common in applied studies to obtain a few cross-loadings with a magnitude up to 0.3 (e.g., Garrido et al., 2020; Tóth-Kiraly, 2017).

Factor Analytic Techniques

Eleven factor analytic techniques were evaluated. All techniques were implemented using maximum likelihood estimation with *Mplus* version 8 via the `MplusAutomation` package (Hallquist & Wiley, 2018) of R software, with two exceptions. First, Markov chain Monte Carlo (MCMC) estimation was used for implementing BSEM (Muthén & Asparouhov, 2012; Muthén & Muthén, 2017). Second, the Saris procedure was implemented with the `lavaan` package because it provides all the information required for the Saris procedure (i.e., MI, SEPC, and statistical power of MI), but *Mplus* does not. The argument `mimic="Mplus"` was used to ensure the maximum similarity with the other techniques. A brief description of each factor analytic technique implementation is provided in the following:

- *Independent clusters model CFA* (ICM-CFA): with major loadings freely estimated and the remaining loadings (i.e., cross-loadings and zero-loadings) fixed to 0.
- *CFA with MI* (CFA_{MI}): after fitting the ICM-CFA, the loading with the highest associated MI, given that the MI was higher than 10.82 ($p \leq 0.001$), was introduced in the model. The process was iteratively repeated until no factor loading fulfilled that criterion.
- *CFA with the Saris procedure* (CFA_S): after fitting the ICM-CFA, the loading with the highest associated MI (higher than 10.82) was introduced in the model if the statistical power of the MI was lower than 0.75 or the SEPC was higher than 0.10. The process was iteratively repeated until no factor loading fulfilled the above criteria. Note that, even though Saris et al. (2009) used an SEPC higher than 0.40 for factor loadings in their study, they recognized that these cutoffs are arbitrary and can be changed for different contexts. Given that cross-loadings with magnitudes lower than 0.40 were generated in the present study, we used 0.10 as the cutoff for factor loadings.
- *EFA with oblique Geomin rotation* (EFA_G): with an epsilon parameter equal to 0.001 for the case of three factors ($K = 3$) and 0.01 for the case of five factors ($K = 5$) following Muthén and Muthén (2017).
- *EFA with oblique target rotation* (EFA_T): with major loadings left as unspecified and the remaining loadings (i.e., cross-loadings and zero-loadings) specified as zero in the partially specified target matrix.

– *EFA with refined target matrix* (RETAM): after fitting the EFA_T, the target matrix was iteratively refined using the RETAM procedure with a complete refinement approach (Lorenzo-Seva & Ferrando, 2020).

– *BSEM*: with non-informative priors for the major loadings and informative priors with small variance (N[0, 0.01]) for the remaining loadings (i.e., cross-loadings and zero-loadings; Guo et al., 2019; Muthén & Asparouhov, 2012). The default options in *Mplus* were used for the MCMC estimation, using a maximum number of iterations of 60,000 for each MCMC chain.

– *ECFA and BCFA with* p*-value criterion* ($ECFA_p$ and $BCFA_p$, respectively): after fitting the EFA_G or BSEM, respectively, a CFA was conducted by freely estimating the resulting statistically significant loadings and fixing to zero the non-significant loadings (α = 0.05).

– *ECFA and BCFA with R-squared criterion* ($ECFA_{R2}$ and $BCFA_{R2}$, respectively): after fitting the EFA_G or BSEM, respectively, a CFA was conducted by freely estimating the resulting relevant loadings and fixing to zero the non-relevant loadings. The R-squared method was used to identify the relevant loadings using thirty values of φ (from 0.70 to 0.99 in steps of 0.01). The BIC was then employed to select the best-fitting solution.

Data Generation

Data were generated from a linear common factor model. Instead of generating a correct model that fits exactly at the population level, which is regarded as unrealistic in applied settings (Lai, 2019; Lorenzo-Seva & Ferrando, 2020), model error was introduced in the data generation process. Namely, the procedure by Cudeck and Browne (1992) was used, which allows to ensure that the minimum of the discrepancy function is achieved at the population model parameter values specified by the researcher, while there is a prespecified degree of model misfit at the population level. Specifically, the population covariance matrix was generated as $\mathbf{\Sigma}^* = \mathbf{\Sigma}(\boldsymbol{\theta}_0) + \Delta_{\text{error}}$, where $\mathbf{\Sigma}(\boldsymbol{\theta}_0)$ is the reproduced variance-covariance matrix and Δ_{error} is a perturbation matrix that is chosen to ensure that $\boldsymbol{\theta}_0 = \arg\min F_{\text{ML}}[\mathbf{\Sigma}^*, \mathbf{\Sigma}(\cdot)]$ and $F_{\text{ML}}[\mathbf{\Sigma}^*, \mathbf{\Sigma}(\boldsymbol{\theta}_0)] = c$. In the present study, the *c* value was chosen in such a way that the root-mean-square error of approximation (RMSEA; Steiger, 1990) was equal to 0.05, which is a commonly encountered value in applied research (Booth & Hughes, 2014). This procedure allows to study the recovery of model parameters ($\boldsymbol{\theta}_0$), even if the model is misspecified. If necessary, the data generation process was repeated until the resulting population correlation matrix $\mathbf{\Sigma}$ was definite positive.

For each of the resulting 128 conditions after combining all the simulation factor levels, one-hundred datasets were generated. Standardized continuous variables were simulated from a multivariate normal distribution using $\mathbf{\Sigma}$ and a sample size of *N* = 300, 650, or 1,000 with the `mvtnorm` package (Genz et al., 2020). The population loading matrices ($\mathbf{\Lambda}$) were randomly generated with the following constraints: (a) each factor was measured by the same number of items (*JK* = 4 or 8) and (b) each factor contained the same number of cross-loadings (*CLK* = 1 or 2). For the remainder of the paper, the population factor loadings with magnitude equal to *ML*, *CL*, and zero will be referred to as major loadings, cross-loadings, and zero-loadings, respectively.

Performance Measures and Data Analysis

The behavior of the factor analytic techniques was evaluated using seven performance measures. First, the convergence rate (CR; i.e., proportion of converged estimations in each condition) and the computation time in seconds (CT) were examined to give an idea of the feasibility and practicality of each technique under each condition. For reference purposes, the simulation was conducted in a desktop computer with 8 3.60-GHz processors and 16GB RAM memory.

Second, parameter estimation accuracy was examined by means of the bias and root-mean-square error (RMSE):

$$Bias(\hat{\boldsymbol{\theta}}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{\boldsymbol{\theta}}_r - \boldsymbol{\theta}) \tag{4}$$

and

$$RMSE(\hat{\boldsymbol{\theta}}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{\boldsymbol{\theta}}_r - \boldsymbol{\theta})^2}, \tag{5}$$

where $\hat{\boldsymbol{\theta}}_r$ is the parameter estimate in replication r, $\boldsymbol{\theta}$ is the generating value of the parameter, and R is the number of converged replications in each condition. Overestimation and underestimation refer to a positive and negative bias, respectively. Equations (4) and (5) were applied separately to major loadings, cross-loadings, zero-loadings, and factor correlations.

In order to better understand the performance similarities of the techniques regarding parameter estimation, an exploratory graph analysis (EGA; Golino & Epskamp, 2017) was conducted on a dataset formed by the bias and RMSE of major loadings, cross-loadings, zero-loadings, and factor correlations of all techniques. Pearson correlations and Gaussian graphical model using graphical LASSO were used for EGA implementation with the `EGAnet` package (Golino & Christensen, 2021). The resulting symmetric network from the EGA was further inspected by means of a multidimensional scaling (MDS) analysis using the `smacof` package (Mair et al., 2021).

An additional analysis was conducted to better summarize the information with the purpose of providing applied guidelines. Namely, a new dependent variable called proportion of successful parameter estimation (PSPE) was defined as the proportion of replications in a specific condition in which a factor analytic technique obtained an $RMSE \leq 0.1$ in major loadings, cross-loadings, zero-loadings, and factor correlations. A non-converged replication was counted as an unsuccessful estimation for the PSPE computation. Thus, the PSPE was intended as a heuristic performance measure that allows to identify not the best method, but rather what methods can be trusted under what conditions. Table 3 illustrates the PSPE under the conditions N = 1,000, JK = 8, FC = 0, and ML = 0 (and the different levels of K, CL, and CLK). For each technique, the average PSPE ($\overline{PSPE}$), minimum PSPE (PSPE_{min}), and maximum PSPE (PSPE_{max}) can be obtained for these specific conditions. A technique was considered as successful under a specific condition if it obtained a $\overline{PSPE} \geq 0.90$ and a $PSPE_{min} \geq 0.70$.

[Please insert Table 3 here]

Beyond parameter estimation, the model fit of the factor analytic techniques was evaluated using the RMSEA and the comparative fit index (CFI; Bentler, 1990):

$$RMSEA = \max\left(\sqrt{\frac{\hat{\delta}_M}{df_M(N-1)}}, 0\right) \quad (6)$$

and

$$CFI = 1 - \frac{\max(\hat{\delta}_M, 0)}{\max(\hat{\delta}_N, \hat{\delta}_M, 0)}, \quad (7)$$

where $\hat{\delta}_M$ and $\hat{\delta}_N$ denote the non-centrality parameter of the specified and null model, respectively, df_M denote the degrees of freedom of the specified model, and N is the sample size. The traditional criteria for model fit evaluation were used in the present study, with RMSEA values lower than 0.08 and 0.05, and CFI values higher than 0.90 and 0.95 as indicators of good and excellent fit, respectively (Hu & Bentler, 1999; Marsh et al., 2004; Yu, 2002). For BSEM, the posterior predictive p-value (PPP; see Asparouhov & Muthén, 2010) was used for the evaluation of model fit. A PPP close to 0.5 indicates good model fit, with values lower than 0.05 regarded as unsatisfactory model fit (Asparouhov & Muthén, 2021).

Lastly, the recovery of the latent factor definitions was evaluated by means of the determinacy of factor scores estimates. Namely, after estimating factor scores using the least squares regression approach (Thurstone, 1935), the squared multiple correlation between each latent factor and the original variables was computed. Both the true reliability (R^2) and the empirical reliability ($\hat{R}^2$) were computed based on either the true structure matrix ($\mathbf{S} = \mathbf{\Lambda\Phi}$) or the estimated structure matrix ($\hat{\mathbf{S}} = \hat{\mathbf{\Lambda}}\hat{\mathbf{\Phi}}$). As suggested by Beauducel (2011), the model-implied correlation matrix ($\hat{\mathbf{\Sigma}}$) was used to compute the determinacy of factor score estimates. Thus, the computation of $\hat{R}^2$ and R^2 is as follows:

$$\hat{R}^2 = \text{diag}(\hat{\mathbf{S}}^{\mathrm{T}}\mathbf{W}\mathbf{L}^{-1})^2 \quad (8)$$

and

$$R^2 = \text{diag}(\mathbf{S}^{\mathrm{T}}\mathbf{W}\mathbf{L}^{-1})^2, \quad (9)$$

where $\mathbf{W} = \hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{S}}$ is the matrix of factor score coefficients and $\mathbf{L}$ is a diagonal matrix of factor score standard deviations, which correspond to the square roots of the diagonal elements of $\mathbf{C} = \mathbf{W}^{\mathrm{T}}\hat{\mathbf{\Sigma}}\mathbf{W}$, which is the latent factors covariance matrix.

To better summarize the results and identify the most relevant main and interaction effects, a univariate ANOVA was conducted for each factor analytic technique and dependent variable. A partial eta squared effect higher than 0.14 was used as a heuristic to label an effect as *relevant* (Cohen, 1988); the effects with a partial eta squared higher than 0.14 were later visually inspected to determine their substantial importance. Thus, the results are presented in a narrative form informed by the ANOVA results but also from a case-by-case inspection of the interactions pointed out as most relevant. The partial eta squared for the different ANOVAs can be found in the Supplementary Material. Moreover, the results of this study are publicly available at https://psychometricmodelling.shinyapps.io/FAcomparison/, and the simulation codes and functions used to implement the factor analytic techniques can be accessed at https://osf.io/vw4xn/ (Nájera et al., 2022a); these functions have been also included in an R package called `wrapFA` (Nájera et al., 2022b), which is a wrapper that allows for factor analysis applications with `lavaan` and `MplusAutomation`. This study was not preregistered.

Results

Feasibility of the Factor Analytic Techniques

All techniques obtained a high overall convergence rate ($CR \geq 0.93$). ICM-CFA and CFA_S obtained a satisfactory convergence rate ($CR > 0.95$) across all conditions. On the contrary, CFA_{MI} and BSEM obtained the lowest overall convergence rates. Specifically, CFA_{MI} had convergence issues with a larger number of items per factor, larger main loadings, and independent factors ($CR = 0.62$), while BSEM had problems with smaller sample sizes, larger number of factors, correlated factors, and smaller main loadings ($CR = 0.27$). Note that BCFA methods are subject to BSEM convergence, so they inherit the same issues. Finally, EFA and ECFA procedures obtained a high overall convergence rate ($0.95 < CR < 0.98$), with lower rates under smaller sample sizes, larger number of factors, smaller number of items per factor, correlated factors, and smaller main loadings ($CR \leq 0.67$), particularly for $ECFA_p$ ($CR = 0.45$).

Regarding computation time, the fastest overall techniques were ICM-CFA and EFA_T ($CT < 3$), and the slowest overall procedure was $BCFA_{R2}$ ($CT = 101$), mainly due to the already high computation time of BSEM ($CT = 86$). All methods became slower as the test length increased, up to an average of 253 seconds for $BCFA_{R2}$ under the condition of 5 factors and 8 indicators per factor.

Similarity of the Factor Analytic Techniques

The EGA suggested four dimensions constituted as (a) ICM-CFA, BSEM, $BCFA_p$, and $BCFA_{R2}$, (b) CFA_{MI} and CFA_S, (c) EFA_T and RETAM, and (d) EFA_G, $ECFA_p$, and $ECFA_{R2}$. Additionally, Figure 1 shows the two-dimensional MDS analysis based on the EGA network. Not surprisingly, all duplets of techniques (i.e., $ECFA_p$ and $ECFA_{R2}$, $BCFA_p$ and $BCFA_{R2}$, CFA_{MI} and CFA_S, EFA_T and RETAM) performed similarly in terms of parameter recovery. Arguably, the most unexpected result is the proximity between ICM-CFA and BSEM, which might be explained by the small variance used for the priors, which overly shrunk the cross-loadings towards zero. The *x-axis* of Figure 1 can be identified as an exploratory-confirmatory continuum with EFA_G and ICM-CFA on each pole, respectively. The five EFA-based methods are placed to the left of the dimension, with CFA and BSEM-based techniques placed to the right. The *y-axis* has a less clear definition, although it is noticeable that the original techniques (i.e., ICM-CFA, BSEM, EFA_G, EFA_T) are placed at the bottom, while the procedures that imply modifications to the original structure are placed at the top. RETAM would be the only misplaced technique according to this interpretation, which can be explained by the great similarity of RETAM and EFA_T outcomes, probably due to the small number of modifications applied by RETAM because of the slight degree of model misspecification. The exploratory-confirmatory continuum found with the multidimensional scaling will be used to provide a sound organization of the techniques' performance in the remainder of the article.

[Please insert Figure 1 here]

Estimation Accuracy of Factor Loadings

Tables 4, 5, and 6 show the marginal means of the bias and RMSE of major loadings, zero-loadings, and cross-loadings, respectively, across the simulation factors. First, all factor analytic techniques obtained an overall accurate estimation of both major loadings and zero-loadings, with a bias and RMSE lower than 0.1. Second, most techniques led to poorer recovery of parameter estimates with smaller sample size, smaller number of items per factor, correlated factors, or smaller main loadings. Third, there were no substantial differences in the accuracy of major loading estimates between the different procedures. If at all, BSEM (as well as ICM-CFA) showed an overestimation tendency of major loadings under correlated factors (*Bias* $\geq$ 0.027), while EFA_G tended to underestimate the magnitude of major loadings under this condition (*Bias* = –0.035). This was also the case for EFA_T and RETAM, although only when the number of factors was higher (*Bias* = –0.030). Fourth, CFA_{MI}, CFA_S, BSEM, BCFA and ECFA methods provided the most accurate zero-loading estimates, with a very low bias. On the other hand, EFA_G, EFA_T, and RETAM showed a slightly worse accuracy of zero-loading estimates. The zero-loading overestimation tendency obtained by the three exploratory techniques was mainly due to the conjunction between correlated factors and small major loadings (*Bias* > 0.022), especially for EFA_G (*Bias* = 0.037).

Second, the estimation of cross-loadings was generally less accurate compared to that of major loadings and zero-loadings. The worst performance occurred again with smaller sample sizes, smaller number of items per factor, correlated factors, or smaller main loadings. Naturally, ICM-CFA obtained the worst results under all conditions. The behavior of BSEM, although slightly better, was like that of ICM-CFA, with a pronounced tendency to underestimate cross-loadings (–0.125 ≤ *Bias* ≤ –0.063) and, subsequently, low accuracy (0.072 ≤ *RMSE* ≤ 0.128). CFA_{MI} and CFA_S, as well as the BCFA methods, improved the respective performance of ICM-CFA and BSEM, leading to similar results: they still showed a slight underestimation tendency (–0.072 ≤ *Bias* ≤ –0.009) and low accuracy under some conditions (0.055 ≤ *RMSE* ≤ 0.133). For the BCFA procedures, the cross-loading underestimation tendency primarily occurred with both small sample sizes and small major loadings (*Bias* = –0.119). Finally, EFA_G, EFA_T, and RETAM provided the most accurate cross-loading estimates (0.042 ≤ *RMSE* ≤ 0.083) and a low bias. The most challenging condition for these techniques, especially for RETAM, was again the conjunction of correlated factors and small major loadings (*RMSE* = 0.092). EFA_G was the only technique that showed an overestimation tendency for cross-loadings, especially with a larger number of items per factor (*Bias* = 0.020) or correlated factors (*Bias* = 0.029). Under these conditions, the ECFA methods alleviated this bias (–0.014 ≤ *Bias* ≤ 0.003) but produced slightly less accurate estimates (0.058 ≤ *RMSE* ≤ 0.095) compared to EFA_G.

[Please insert Table 4 here]

[Please insert Table 5 here]

[Please insert Table 6 here]

Estimation Accuracy of Factor Correlations

Table 7 displays the marginal means of the bias and RMSE of factor correlations across the simulation factors. ICM-CFA showed the greatest overestimation tendency (0.066 ≤ *Bias* ≤ 0.152), followed closely by BSEM (0.051 ≤ *Bias* ≤ 0.105). Both techniques overestimated factor correlations especially with a smaller number of factors, smaller number of items per factor, smaller main loadings, or larger magnitude and number of cross-loadings. CFA_{MI}, CFA_S, $BCFA_p$, and $BCFA_{R2}$ provided less biased factor correlation estimates, showing only a slight overestimation tendency, particularly under the abovementioned conditions. On the other hand, EFA_G obtained the most underestimated factor correlation estimates overall, which was mainly caused by the condition of correlated factors (*Bias* = –0.115), since it did not show any bias with independent factors (*Bias* = 0.001). Provided that the factors were correlated, a smaller magnitude of major loadings led to a greater underestimation tendency of factor correlation for EFA_G (*Bias* = –0.170). Larger cross-loadings also contributed to a greater underestimation tendency. A similar tendency was observed for the ECFA procedures, although they showed a much milder underestimation of factor correlations when the factors were indeed correlated (–0.046 ≤ *Bias* ≤ –0.043). Finally, EFA_T and RETAM showed an overestimation tendency with independent factors (0.057 ≤ *Bias* ≤ 0.059) and an underestimation tendency with correlated factors (*Bias* = –0.035). The latter was caused by the interaction between correlated factors and small major loadings (*Bias* = –0.087), since these techniques showed no bias with correlated factors and large major loadings (*Bias* = 0.008).

The accuracy of factor correlation estimates is mostly congruent with the bias results, since ICM-CFA, BSEM, and EFA_G, which were the most biased techniques, provided the least accurate factor correlation estimates, with an overall *RMSE* higher than 0.097. Again, ICM-CFA and BSEM were particularly affected by smaller sample sizes, smaller number of factors, smaller number of items per factor, smaller main loadings, or larger magnitude and number of cross-loadings, while the bad overall performance for EFA_G was mostly due to the interaction between correlated factors and small major loadings (*RMSE* = 0.195). In fact, EFA_G obtained the most accurate factor correlations with independent factors (*RMSE* = 0.061). The remaining techniques also provided overall accurate factor correlations (overall *RMSE* ≤ 0.085), with EFA_T and RETAM performing particularly well under the presence of both correlated factors and large major loadings (*RMSE* = 0.049).

[Please insert Table 7 here]

Proportion of Successful Parameter Estimation

Figure 2 shows a decision tree based on the PSPE results. The decision tree was constructed such that it contained the minimum possible number of terminations. A termination was achieved when either: (a) at least one factor analytic technique obtained a successful parameter recovery (i.e., $\overline{PSPE} \geq 0.90$ and $PSPE_{min} \geq 0.70$), (b) no technique obtained a successful parameter recovery but there was at least one method with $PSPE_{max} \geq 0.50$, or (c) no technique obtained a $PSPE_{max} \geq 0.50$ and thus no method could be recommended under that condition. Accordingly, the higher is a simulation factor in the decision tree, the greater its effect on the methods functioning. The simulation factors *CL* and *CLK* do not appear in the tree for parsimony reasons due to their relatively low relevance, which did not substantively change the outcome of the decision tree. The sample size of *N* = 1000 was not included in the figure because the results were congruent with those of *N* = 650; that is, the best performing techniques in each condition were the same for both sample sizes. Researchers should expect the PSPE to be even higher for these techniques under *N* = 1000.

A notable finding is that, provided that the major loadings are high, a successful parameter estimation can be expected either by EFA_G if the factors are uncorrelated or by

EFA_T and RETAM if the factors are correlated. If the major loadings are low, the decision tree becomes more complex. In general, EFA_G is still the preferred method under independent factors, obtaining the highest PSPE in all conditions. Note, however, that the PSPE of EFA_G is still very low under some scenarios, such as smaller sample sizes, smaller number of items per factor, and larger number of factors ($\overline{PSPE}$ = 0.54). Whenever factors are correlated, EFA_T obtained the most satisfactory results with a smaller number of attributes, while BSEM or $BCFA_p$ performed the best with a larger number of attributes, even though the PSPE under a smaller number of items per factor was insufficient. In this vein, note that there are some conditions under which no technique achieved a $PSPE_{max} \geq 0.50$ (e.g., smaller main loadings, correlated factors, and smaller sample size).

[Please insert Figure 2 here]

Model Fit

Table 8 shows the marginal means of CFI and RMSE across the simulation factors. All techniques achieved a higher CFI under a smaller number of factors, a smaller number of items per factors, correlated factors, and larger main loadings, as well as with a larger magnitude and number of cross-loadings (except for ICM-CFA). As expected, ICM-CFA obtained consistently the worst fit across all conditions ($0.800 \leq CFI \leq 0.901$; $0.056 \leq RMSEA \leq 0.072$). It should be noted that, following the traditional cutoff values, ICM-CFA would have been categorized as having bad fit according to CFI, but good fit according to RMSEA. CFA_{MI} and CFA_S improved model fit in comparison to ICM-CFA, with an overall CFI higher than 0.895. ECFA and BCFA methods obtained a similar, slightly higher overall CFI values ($0.903 \leq CFI \leq 0.908$), only surpassed by the EFA methods (CFI = 0.913). Unlike CFI, RMSEA values were almost identical throughout all conditions and techniques (except for ICM-CFA), being the RMSEA almost identical to the one used for the data generation process (i.e., *RMSEA* = 0.05). Lastly, BSEM obtained unsatisfactory model fit across all conditions (*PPP* ≤ 0.014; see Table 8).

[Please insert Table 8 here]

Determinacy of Factor Score Estimates

Table 9 shows the marginal means of the determinacy of factor score estimates across the simulation factors. Both empirical and true reliabilities are provided. Overall, the empirical reliability was very similar for all factor analytic techniques (0.786 ≥ $\hat{R}^2$ ≥ 0.804). Reliability was overestimated for all procedures ($\hat{R}^2 - R^2$ ≥ 0.033), especially for the exploratory methods (i.e., EFA_G, EFA_T, RETAM; $\hat{R}^2 - R^2$ ≥ 0.063).

Regarding the true reliability, all factor analytic techniques obtained a very similar overall determinacy (0.735 ≤ R^2 ≤ 0.760), with lower values under smaller main loadings, smaller number of items per factor, or independent factors. If at all, the exploratory methods obtained comparatively slightly lower true reliabilities under smaller sample sizes (R^2 ≤ 0.723), larger number of factors (R^2 ≤ 0.732), correlated factors (R^2 ≤ 0.763), and smaller main loadings (R^2 ≤ 0.636).

[Please insert Table 9 here]

Real Data Illustration

The functioning of the eleven factor analytic techniques is illustrated by using real data. The employed dataset was first used in Roskam et al. (2015) and consists of the responses of 2,532 participants to the French version of the Personality Inventory for DSM-5 (PID-5), originally developed by Krueger et al. (2012).[3] The PID-5 is formed by 220 items assessed with a 4-point Likert-type scale. The items are grouped into 25 facets, which are in turn grouped into 5 domains: Negative affect, Detachment, Antagonism, Disinhibition, and Psychoticism. The relationship between the domains, facets, and items can be consulted in

[3] This dataset is publicly available at https://doi.org/10.1371/journal.pone.0133413.s001.

the Supplementary Material. Many studies have been already conducted on the PID-5, revealing good psychometric properties of the scale (see Somma et al., 2019; Sorrel et al., 2021; Watters & Bagby, 2018). Namely, Somma et al. (2019) conducted the most extensive meta-analytic review of validation studies concerning the PID-5, with 23 published articles and a total sample size of N = 24,240. The final EFA solution they found with the pooled correlation matrix (using the 25 facets and the 5 domains) will be here considered as a baseline for comparing the behavior of the eleven factor analytic techniques (see Table 5 in Somma et al., 2019). The root-mean-square deviation (RMSD) was used to compare the solution provided for each technique compared to the one found by Somma et al. (2019). The RMSD is defined as the RMSE in Equation (5), although using the parameter estimates found by Somma et al. (2019) instead of the generating parameters (since this is a real application and the true parameters are unknown). For those methods that require a prespecified structure (i.e., ICM-CFA, BSEM, EFA_T), the expected domain for each facet as identified in Somma et al. (2019) was considered as a target loading. It is noteworthy that Somma et al. (2019) used an oblique target rotation procedure and the weighted least square estimator in their analysis. Considering that facets are continuous variables, our analyses were performed using the robust maximum likelihood estimator, which has been employed more commonly for the PID-5 evaluation (e.g., Sorrel et al., 2021; Thimm et al., 2017), including the original study by Krueger et al. (2012).

The different techniques were implemented as indicated in the Method section with a few exceptions. First, CFA_{MI} did not converge after introducing 42 modifications; just with the purpose of obtaining some results for the comparison, the outcome from iteration 41 was used in these analyses. Note that, in line with the simulation study, only loadings were allowed to become free (no residual covariances were considered). Second, $ECFA_p$ did not achieve convergence with the criterion of α = 0.05. This is mainly because of the great number of significant free loadings (86 loadings out of 125) caused by the large sample size (*N* = 2,532). Forty-six of these loadings had an absolute magnitude lower than 0.10. In order to maintain the $ECFA_P$ method in the illustration analyses, the nominal level was corrected by Bonferroni to reduce the number of free parameters (Cudeck & O'Dell, 1994).

According to the results found by Somma et al. (2019), the conditions of the real data illustration were the following: *N* = 2,532, *K* = 5, *JK* = 5, *FC* = 0.14, *ML* = 0.56, *CL* = 0.24, and *CLK* = 9.20. *FC* was computed as the average absolute factor correlations, *ML* as the average absolute magnitude of the highest loading for each facet, *CL* as the average magnitude of the loadings with an absolute value higher than 0.15 (excluding the highest loading for each item), and *CLK* as the number of loadings with an absolute value higher than 0.15 (excluding the highest loading for each item) divided by the number of factors (i.e., *K* = 5). All values were close to or in between the levels employed in the simulation study, except for the sample size (much larger, expectedly benefiting the performance of the techniques) and the number of cross-loadings per factor (much larger, expectedly disrupting the performance of the techniques, especially the more confirmatory ones). According to the simulation study results, EFA_G is expected to perform the best under these conditions (considering *ML* = 0.5, *FC* = 0, *N* = 1000, *JK* = 4, and *K* = 5 in Figure 2).

To evaluate the stability and replicability of the factor structure found by each technique, we conducted a nonparametric bootstrap resampling procedure as explained in Christensen and Golino (2021). Namely, for each of 100 replications, 2,532 participants were randomly drawn from the original database with a replacement. Then, each factor analytic technique was applied to the resampled data and, once the 100 replications were done, the congruent coefficient (CC) of the factor loading matrix between pairs of replicates was computed for each factor. The average CC across the 100 replications is used as a measure of the stability of the factor structure.

The performance of the techniques is summarized in Table 10. BSEM, $BCFA_p$, and $BCFA_{R2}$ are not presented because BSEM did not achieve convergence. First, ICM-CFA obtained the highest major loadings (*ML* = 0.621) and factor correlations (*FC* = 0.621), with an unsatisfactory model fit (*CFI* = 0.689; *RMSEA* = 0.112). The highest stability obtained by ICM-CFA across replications (*CC* = 0.999) is inherent to the fact that the hypothesized structure is not modified in this procedure (EFA_T, which is the other technique based on an unmodifiable structure, shows a similar result in this regard). Unexpectedly, CFA_{MI} and CFA_S did not provide similar results, mainly due to CFA_{MI} including 42 modifications in the model and CFA_S including only 20 modifications. For instance, CFA_{MI} obtained larger major loadings (*ML* = 0.633), cross-loadings (*CL* = 0.144), and factor correlations (*FC* = 0.364) compared to CFA_S (*ML* = 0.566; *CL* = 0.068; *FC* = 0.322). Additionally, the smaller number of modifications included in CFA_S resulted in a worse fit (*CFI* = 0.835; *RMSEA* = 0.092). Moreover, both procedures obtained the lowest stability across replications (*CC* < 0.893), which is consistent with the literature that argues that these procedures lack replicability due to the problem of capitalization on chance (Browne, 2001; MacCallum et al., 1992). On the contrary, the EFA-related techniques provided similar average major loadings, cross-loadings, and factor correlations. The pattern of results of EFA and ICM-CFA supports the simulation study findings where exploratory techniques showed: (a) lower factor correlation estimates, (b) lower factor score determinacies (e.g., $\hat{R}^2$ = 0.871 for ICM-CFA and 0.833 for EFA_G), and (c) better fit (e.g., CFI = 0.689 for ICM-CFA and 0.883 for EFA_G). EFA_G also showed a high degree of stability across replications (*CC* = 0.964), followed closely by $ECFA_{R2}$ (*CC* = 0.958). RETAM and $ECFA_p$ showed a slightly lower congruence (*CC* = 0.942 and 0.933, respectively). The latter obtained a *CC* > 0.939 for all domains except for Disinhibition, which achieved a *CC* = 0.804.

[Please insert Table 10 here]

EFA_T obtained the most similar structure compared to that of Somma et al. (2019), with an *RMSD* = 0.088 for factor loadings and *RMSD* = 0.173 for factor correlations, followed by RETAM (*RMSD* = 0.109 and 0.202, respectively). EFA_G, $ECFA_p$, and $ECFA_{R2}$ also obtained similar estimates compared to Somma et al. (2019; $0.177 \leq RMSD \leq 0.181$ and $0.205 \leq RMSD \leq 0.221$ for factor loading and correlation estimates, respectively). The higher similarity of EFA_T over EFA_G is probably because a target rotation procedure was also used in Somma et al. (2019). However, according to the simulation study results, under these specific conditions EFA_G might be considered particularly, since it provided a more accurate estimation of the generating model under similar scenarios. In any case, the interpretation of the final EFA_T and EFA_G models did not greatly differ; as shown in Table 11, they coincided in 87.2% of the loadings when categorizing them as relevant ($\lambda \geq 0.3$) and irrelevant ($\lambda < 0.3$). Moreover, the definition of all domains remained similar in both solutions, although Detachment and Disinhibition were less clearly defined for EFA_G. The facets from the former cross-loaded on Negative affectivity, a tendency that could be already found in Somma et al. (2019) and that has been encountered in other scales beyond PID-5 (e.g., Sorrel et al., 2022). Disinhibition obtained the lowest main loadings of all domains for both EFA_G and EFA_T; this is congruent with other studies that have found that Disinhibition tends to blend with other constructs (Oltmanns & Widiger, 2020; Sorrel et al., 2022). This was especially true for EFA_G, where some of the Disinhibition facets primarily loaded on Psychoticism. This also explains why the stability of this domain was the lowest not only for EFA_G (*CC* = 0.875), but for other techniques such as $ECFA_p$.

[Please insert Table 11 here]

Discussion

The increasing number of factor analytic techniques that have been developed in the last years might not have an actual impact on applied settings if practitioners cannot have a clear idea of what technique should be preferred under what circumstances. The present paper had the main purpose of providing such guidelines based on the results of an exhaustive simulations study. The following research questions were addressed:

(Q1) *How can the factor analytic techniques be grouped or categorized according to their functioning across conditions when the model (i.e., initial internal structure) is slightly misspecified?* The theoretical classification of factor analytic techniques along the confirmatory-exploratory continuum was supported by the empirical evaluation of their parameter estimates. In line with the previous literature (Guo et al, 2019; Xiao et al., 2019), the more confirmatory methods (i.e., ICM-CFA, CFA_{MI}, CFA_{S}, BSEM, BCFA) tended to underestimate cross-loadings, overestimate factor correlations, and show worse model fit. The functioning of ICM-CFA and BSEM was more similar than expected, which might be caused by the small variance used for the non-target loadings priors. On the other hand, the more exploratory methods (i.e., EFA_{G}, ECFA, EFA_{T}, RETAM) showed better model fit and an overall more accurate parameter estimation.

(Q2) *When the baseline model is theoretically constrained (i.e., ICM-CFA) but slightly misspecified, does sequential model modification substantially improve the recovery of the internal structure of a test?* CFA_{MI} and CFA_{S} obtained more accurate parameter estimates than ICM-CFA, especially a milder overestimation of factor correlations due to the inclusion of cross-loadings. The performance of these sequential model modifications was more satisfactory than what had been previously found in the literature (Whittaker, 2012; Yuan & Liu, 2021), probably because the initial models were only slightly misspecified in the present study. Even though a small number of misspecifications is expected in those research fields with a well-developed theoretical corpus, more novel topics will often contain a certain degree of uncertainty regarding the specification of the internal structure. The performance of CFA_{MI} and CFA_{S} will depend on the degree of model misspecification present in the initial model formulation. In line with MacCallum et al. (1992), their use should be reserved to situations where there is strong a priori knowledge, the sample size is large enough to provide consistent results, and the number of suggested modifications is low. Otherwise, the temptation of making model adjustments to improve model fit might lead to the capitalization on chance problem, as shown in the real data example by the low stability of their solutions. Moreover, the use of arbitrary cutoff points by the Saris procedure is an additional practical burden that might hinder its applicability, and practitioners should be aware that, as a chi-square index, the optimal cutoff used to identify relevant model parameters (e.g., MI > 10.82) is highly dependent on sample size (see, e.g., Whittaker, 2012).

(Q3) *When the baseline model is unconstrained but theoretically guided (e.g., BSEM, EFA_{T}) and slightly misspecified, does adding model constraints (i.e., BCFA) or refining the initial internal structure (i.e., RETAM) provide substantial advantages?* The BCFA techniques provided more accurate parameter estimates than BSEM, showing less underestimated cross-loadings and, subsequently, less overestimated factor correlations. These procedures showed an overall satisfactory model fit, while the BSEM obtained a consistently poor model fit according to the PPP. This latter result should be interpreted with caution, since the PPP might be overly conservative (Asparouhov & Muthén, 2010, 2021). The development and further study of the Bayesian adaptations of commonly known fit indices (e.g., RMSEA, CFI) might facilitate the comparison of frequentist and Bayesian model fit (Asparouhov & Muthén, 2021). On the other hand, RETAM performed very similarly to EFA_{T}. This might be due to EFA_{T} being able to accommodate slightly misspecified models due to its unrestricted nature. RETAM is expected to perform better than EFA_T under those conditions in which the distance between the original target matrix and the generating model is larger (Lorenzo-Seva & Ferrando, 2020).

(Q4) *When the baseline model is unconstrained and not guided by the theory (i.e., EFA_G), does adding model constraints (i.e., ECFA) provide substantial advantages?*

Partially. The ECFA procedures obtained slightly less overestimated zero-loadings and, subsequently, less underestimated factor correlations compared to EFA_G. This was particularly true when the factors were indeed correlated (*FC* = 0.5), because EFA_G showed a large underestimation tendency (*Bias* = –0.177), while $ECFA_p$ and $ECFA_{R2}$ reduced the bias (*Bias* ≥ –0.047). Accordingly, ECFA might be preferred over EFA_G with clearly correlated factors. On the other hand, EFA_G provided slightly more accurate cross-loadings overall. Moreover, the three techniques obtained very similar CFI and RMSEA values. This reflects that ECFA can be applied without the risk of showing inflated fit indices.

(Q5) *Do the hypothesis-based techniques (i.e., by fixing, targeting, or setting priors for expected zeros) substantially improve the performance of the theoretically blind techniques (i.e., EFA_G, ECFA) when the model is only slightly misspecified?*

Partially. EFA_G obtained the most accurate parameter estimates when the factors were independent (*FC* = 0). However, it overestimated zero-loadings and underestimated factor correlations when the factors were correlated (*FC* = 0.5). Under these conditions, EFA_T and RETAM generally provided the best performance. This is unfortunate for EFA_G since many instruments are designed for the measurement of correlated psychological dimensions. To further explore the performance of data-driven techniques (i.e., EFA_G, $ECFA_p$, $ECFA_{R2}$) with correlated factors, Figure 3 extends the decision tree based on the PSPE by focusing only on these procedures under *ML* = 0.7 and *FC* = 0.5. Here, it is shown that EFA_G maintained a reasonable performance for *N* = 1000 and *K* = 5, but $ECFA_p$ might be generally a better option, particularly with lower sample sizes (*N* = 300).

[Please insert Figure 3 here]

(Q6) *Is there a clear correspondence between parameter estimation accuracy, model fit indicators, and determinacy of factor score estimates?* The simulation study showed that this is not the case. As previously suggested, good model fit can be achieved with inaccurate parameter estimates (MacCallum et al., 1992). In this vein, ICM-CFA obtained an $RMSEA \leq 0.08$ across all conditions. However, we have also observed the opposite: bad model fit achieved with accurate parameter estimates. Thus, a $CFI \leq 0.90$ was observed for all techniques under several conditions (e.g., *FC* = 0), despite the good parameter estimation accuracy obtained by some of the techniques. These findings support the repeated notion that the cutoffs for model fit indices should not be blindly taken as golden rules, but contextualized (Hu & Bentler, 1999; Marsh et al., 2004; Yu, 2002). Furthermore, the empirical reliability of the factor score estimates was virtually identical for all factor analytic techniques despite the differences in parameter estimation, which is congruent with previous findings (Booth & Hughes, 2014). The determinacy of factor score estimates was, however, overestimated, especially for the exploratory procedures. In fact, ICM-CFA obtained a higher true reliability than EFA_G, EFA_T, and RETAM. Thus, even though ICM-CFA does not properly recover cross-loadings and factor correlations under slightly misspecified models, it could be still used to compute factor scores as long as the theoretical structure of the test can be clearly specified. Otherwise, a largely misspecified model would disrupt factor score estimates to a greater extent.

Four procedures were proposed in the present study: $ECFA_p$, $ECFA_{R2}$, $BCFA_p$, and $BCFA_{R2}$. Based on Figure 1, these methods are located towards the center of the exploratory-confirmatory continuum. Thus, compared to EFA_G, ECFA improved the recovery of major loadings, zero-loadings, and factor correlations (under the presence of correlated factors) at the cost of obtaining more inaccurate cross-loading estimates. On the other hand, even though BCFA is technically a more restrictive model than BSEM, its results were in fact less similar to that of the most restrictive technique (i.e., ICM-CFA). Moreover, BCFA consistently outperformed BSEM in terms of parameter estimation accuracy, and generally obtained accurate estimates for major loadings, zero-loadings, and factor correlations, at the expense of underestimating cross-loadings. Thus, whenever BSEM is used in scale validation studies, BCFA is also recommended to compare both solutions and evaluate the discrepancies in parameter estimates. Nonetheless, the high computational cost and potential convergence issues might hinder the applicability of these techniques in real settings. In line with the guidelines provided by Bandalos and Finney (2019), a series of steps should be followed when conducting and reporting the results from ECFA or BCFA: (a) clearly present the results from the original EFA or BSEM analysis, (b) compare the results from the original model (EFA or BSEM) with that of the final solution (ECFA or BSEM), (c) transparently notify the unexpected findings (e.g., cross-loadings) in both the original and final solution and whether they should be considered from a substantial point of view or not, (d) try to replicate the results using either cross-validation (if a second sample is available) or the bootstrapping procedure explained in the Real Data Illustration section.

Based on the above, a series of practical recommendations can be stated. First, given that confirmatory and exploratory techniques have opposed strengths and weaknesses, a safe practice when conducting a validation study would consist in analyzing the data with one of each and then evaluate to what extent the results of both approaches are congruent (Garrido et al., 2020; Schmitt et al., 2018). Of course, the simulation study results imply that more credit should be given to EFA_G whenever factors are approximately independent, and to EFA_T and RETAM whenever the factors are clearly correlated. In this vein, the decision tree shown in Figure 2 can give an idea of what technique is expected to perform the best under a specific set of conditions. Moreover, it should be noted that, under certain conditions, no factor analytic technique is expected to provide reliable results. To avoid these scenarios, researchers can attempt to create larger tests and have larger sample sizes for validation studies.

It could be argued that the reason why we do not recommend the use of ICM-CFA to explore the internal structure of questionnaires based on the findings of the simulation study is because we did not include any simulation condition without cross-loadings. In other words, the fact that ICM-CFA performed poorly in the simulation study (were all conditions included some degree of cross-loadings) could be regarded as a trivial finding. We believe that this is not the case, since one of the relevant results from this simulation study (as well as some previous research; e.g., Asparouhov & Muthén, 2009) is that ICM-CFA provides very biased parameter estimates even when the model is just slightly misspecified. That is, only one or two cross-loadings per factor were included in the simulation study; this might be even considered as a benevolent scenario for the confirmatory methods, since the majority of empirical applications in psychological sciences often encounter more and larger cross-loadings (Booth & Hughes, 2014; Hopwood & Donnellan, 2010). As a final note, please note that if we had generated items without cross-loadings, then the true, simple structure ICM-CFA model would probably have performed much better. Similarly, BSEM is expected to show a very good performance when all the priors are correctly specified (Xiao et al., 2019). In summary, even though ICM-CFA and BSEM are arguably the most appropriate techniques when the hypothesized structure is indeed correct, it is imperative to critically consider whether this is a plausible scenario in most real settings (Asparouhov & Muthén, 2009; Booth & Hughes, 2014; Hopwood & Donnellan, 2010; MacCallum et al., 2012; Marsh et al., 2014; Xiao et al., 2019). Testing the replicability of such a structure in multiple samples will be essential in any case (Franco-Martínez et al., 2022).

On another note, we have focused on explaining the main effects of the simulation factors, along with a few relevant interaction effects, due to the impossibility of reflecting the whole complexity of the simulation study. Thus, in addition to Figure 2, which summarizes in a practical fashion the preferred factor analytic techniques under different conditions, we have created a Shiny app (Chang et al., 2021) in which the interested reader can interactively explore the performance of each technique under any condition, visualizing the distribution of the different dependent variables as a function of various interaction effects between the independent variables.[4] The ANOVA results available in the Supplementary Material can serve as a guide to explore the desired effects using the app. Furthermore, in order to increase the applied repercussion of the research, all the R codes corresponding to the factor analysis functions, simulation study and data analysis were made available at https://osf.io/vw4xn/.

The present study is not without limitations. First, standardized continuous variables had been employed in the simulation study. This setting relates to those instruments in which facets, rather than items, are used as indicators. Some well-known examples of these instruments are the PID-5 (Krueger et al., 2012), Five-Factor Personality Inventory for ICD-11 (Sorrel et al., 2022), or facet-level analyses of the NEO inventories (Lui et al., 2020). Although the results of the present study are mostly congruent with those obtained by Liang et al. (2021) with categorical indicators, further research is needed to evaluate whether our conclusions hold when normality cannot be assumed. Second, and related to this, we only evaluated the performance of the different factor analytic techniques using maximum

[4] https://psychometricmodelling.shinyapps.io/FAcomparison/

likelihood estimation which, even though it is one of the recommended estimation methods when data follow a multivariate normal distribution (Goretzko et al., 2019), it is also more prone to convergence issues or to fail to identify small factors than other estimation procedures (Briggs & MacCallum, 2003). A good practice in real settings would be to evaluate the consistency of the factor solution between different estimation methods, such as unweighted least squares. Third, the identification of correlated residuals was not considered in the present study. The detection of relevant correlated residuals, in conjunction with the cross-loadings here explored, should be further studied. Fourth, we only explored the performance of EFA with Geomin rotation, although it should be highlighted again that other rotation procedures might lead to different results (Browne, 2001; Hakstian, 1971; Hakstian & Abell, 1974; Sass & Schmitt, 2010; Schmitt & Sass, 2011). Recently, Nguyen and Waller (2022) found that Geomin is more prone to converge to local minima than other rotation procedures. For the sake of the present investigation, and following the suggestion from one of the reviewers, we replicated the simulation study using EFA with Oblimin rotation and found that the results did not substantially differ from those of EFA with Geomin. Compared to Geomin, Oblimin tended to provide more biased parameter estimated under the presence of independent factors (i.e., EFA with Geomin is still the recommended technique for these scenarios) and provided less biased estimates with correlated factors, although it did not outperform EFA with Target rotation or the RETAM procedure in these settings. In any case, practitioners should be aware that the choice of the rotation procedure is nontrivial, and that different rotation methods are better suited to different settings. Accordingly, further research should be conducted regarding the exploration of the performance of ECFA with different rotation procedures. Fifth, BSEM underperformed in comparison with previous studies (Guo et al., 2019; Xiao et al., 2019; Wei et al., 2022) due to the misspecified priors with small variance. Using larger variance for the BSEM priors might reduce the bias of the estimates, at the cost of increasing convergence issues (MacCallum et al., 2012; Muthén & Asparouhov, 2012). Note that BSEM already had the highest nonconvergence rates among all techniques. It is noteworthy that BSEM showed the lowest convergence rate with small sample sizes, where Bayesian methods are expected to perform better than frequentists (Muthén & Asparouhov, 2012). These findings enhance the importance of conducting sensitivity analyses for prior selection, even though it might be a practical burden (MacCallum et al., 2012; Rindskopf, 2012). Moreover, the model fit evaluation of BSEM is still to be investigated to achieve an easy interpretation and known type I error rate (Asparouhov & Muthén, 2021). In this vein, Bayesian adaptations of those indices have been recently proposed, but further research is still required to understand their functioning (Asparouhov & Muthén, 2021). Sixth, only RMSEA and CFI were here calculated to evaluate model fit, but other indices such as the Tucker-Lewis index (TLI; Tucker & Lewis, 1980) or the standardized root mean square residual (SRMR; Jöreskog & Sörbom, 1988) are also worth exploring. Finally, a few techniques could not be included in the simulation study due to computational burden. This is the case of regularized structural equation modeling (Jacobucci et al., 2016) and penalized likelihood structural equation modeling (Huang et al., 2017). Moreover, the inclusion of different BSEM priors would be also valuable according to the importance of sensitivity analyses. However, it is expected that the results would not substantially differ if different variances were used for the priors (Liang et al., 2020; Xiao et al., 2019). Regarding the ECFA procedures, further research could focus on evaluating their performance with slightly different criteria for the identification of the relevant factor loadings. For instance, it could be explored whether using a multiple comparison correction (e.g., Bonferroni) systematically improves the performance of the $ECFA_p$ or, on the contrary, leads to overly simple structures. Similarly, a more liberal criterion than the BIC, such as the Akaike information criterion (Akaike, 1974), could be used for the selection of the optimal φ

in the $ECFA_{R2}$, noticing that it would lead to more complex final models. Other approaches, such as the Promin criterion (Lorenzo-Seva, 1999) or the SLiD criterion (Garcia-Garzon et al., 2018) could be also explored as methods to specify the optimal cutoffs to identify relevant loadings.

Lastly, it should be highlighted again that an appropriate implementation of all factor analytic techniques goes beyond methodological knowledge. Since all techniques (but ICM-CFA) are prone to capitalization on chance, given that they have some degree of data-driven component (MacCallum et al., 1992), substantial interpretation is crucial to increase the likelihood of a model being replicable and generalizable to other samples and situations; this should be complemented, whenever possible, with a cross-validation study to empirically evaluate the consistency of the findings. In other words, even though a high emphasis has been put into methodological advances in the last decades, with numerous new techniques and variations, theoretical interpretation and applicability of the models should not be put aside (Ferrando, 2021). The inclusion of a parameter that cannot be understood nor defended in substantive terms is a problem that should not be blindly accepted regardless of what factor analytic technique has been implemented.

References

Akaike, H. (1974). A new look at the statistical identification model. *IEEE Trans. Automated Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Arens, A. K., & Morin, A. J. S. (2016). Examination of the structure and grade-related differentiation of multidimensional self-concept instruments for children using ESEM. *The Journal of Experimental Education*, *84*(2), 330–355. https://doi.org/10.1080/00220973.2014.999187

Asparouhov, T., & Muthén, B. (2021). Advances in Bayesian model fit evaluation for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(1), 1–14. https://doi.org/10.1080/10705511.2020.1764360

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis using Mplus: Technical implementation.* Technical Report. Version 3. http://statmodel.com/download/Bayes3.pdf

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Bandalos, D. L., & Finney, S. J. (2019). Factor analysis: Exploratory and confirmatory. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (2nd ed., pp. 98–122). Routledge.

Beauducel, A. (2011). Indeterminacy of factor score estimates in slightly misspecified confirmatory factor models. *Journal of Moderns Applied Statistical Methods*, *10*(2), 583–598. https://doi.org/10.22237/jmasm/1320120900

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Booth, T., & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, *21*(3), 260–271. https://doi.org/10.1177/1073191114528029

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

Chang, W., Cheng, J., Allaire, J. J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2021). *shiny: Web Application Framework for R. R package version 1.6.0*. https://CRAN.R-project.org/package=shiny

Chou, C.-P., & Bentler, P. M. (1993). Invariant standardized estimated parameter change for model modification in covariance structure analysis. *Multivariate Behavioral Research*, *28*(1), 97–110. https://doi.org/10.1207/s15327906mbr2801_6

Christensen, A. P., & Golino, H. (2021). Estimating the stability of psychological dimensions via boostrap exploratory graph analysis: A Monte Carlo simulation and tutorial. *Psych*, *3*, 479–500. https://doi.org/10.3390/psych3030032

Clarkson, D. B., & Jennrich, R. I. (1988). Quartic rotation criteria and algorithms. *Psychometrika*, *53*(2), 251–259. https://doi.org/10.1007/BF02294136

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, *2nd Edn*. Erlbaum.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment*, *4*(1), 5–13.

Cudeck, R., & Browne, M. W. (1992). Constructing a covariance matrix that yields a specified minimizer and a specified minimum discrepancy function value. *Psychometrika*, *57*(3), 357–369. https://doi.org/10.1007/BF02295424

Cudeck, R., & O’Dell, L. L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*, *115*(3), 475–487. https://doi.org/10.1037/0033-2909.115.3.475

de la Torre, J., & Chiu C.-Y. (2016). A general method of empirical Q-matrix validation. *Psychometrika*, *81*(2), 253–273. https://doi.org/10.1007/s11336-015-9467-8

DiStefano, C., Liu, J., & Burgess, Y. (2017). Investigating the structure of the Pediatric Symptoms Checklist in the preschool setting. *Journal of Psychoeducational Assessment*, *35*(5), 494–505. https://doi.org/10.1177/0734282916647648

Ferrando, P. J. (2021). Seven decades of factor analysis: From Yela to the present day. *Psicothema*, *33*(3), 378–385. https://doi.org/10.7334/psicothema2021.24

Fong, T. C. T., Chan, J. S. M., Chan, C. L. W., Ho, R. T. H., Ziea, E. T. C., Wong, V. C. W., Ng, B. F. L., & Ng, S. M. (2015). Psychometric properties of the Chalder Fatigue Scale revisited: An exploratory structural equation modeling approach. *Quality of Life Research*, *24*, 2273–2278. https://doi.org/10.1007/s11136-015-0944-4

Fong, T. C. T., & Ho, R. T. H. (2013). Factor analyses of the Hospital Anxiety and Depression Scale: A Bayesian structural equation modeling approach. *Quality of Life Research*, *22*, 2857–2863. https://doi.org/10.1007/s11136-013-0429-2

Franco-Martínez, A., Alvarado, J. M., & Sorrel, M. A. (2022). Range restriction affects factor analysis: Normality, estimation, fit, loadings, and reliability. *Educational and Psychological Measurement*. https://doi.org/10.1177/00131644221081867

Garcia-Garzon, E., Abad, F. J., & Garrido, L. E. (2018). Improving bi-factor exploratory modelling: Empirical target rotation based on loading differences. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *15*(2), 45–55. https://doi.org/10.1027/1614-2241/a000163

Garrido, L. E., Barrada, J. R., Aguasvivas, J. A., Martínez-Molina, A., Arias, V. B., Golino, H. F., Legaz, E., Ferrís, G., & Rojo-Moreno, L. (2020). Is small still beautiful for the Strengths and Difficulties Questionnaire? Novel findings using exploratory structural equation modeling. *Assessment*, *27*(6), 1349–1367. https://doi.org/10.1177/1073191118780461

Genz, A., Bretz, F., Miwa, T., & Mi, X. (2020). *mvtnorm: Multivariate normal and t distributions*. *R package version 1.1-1*. https://cran.r-project.org/web/packages/mvtnorm

Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2013). Further insights on the French WISC-IV factor structure through Bayesian structural equation modeling. *Psychological Assessment*, *25*(2), 496–508. https://doi.org/10.1037/a0030676

Golino, H., & Christensen, A. (2021). *EGAnet: Exploratory graph analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics*. *R package version 0.9.8*. https://cran.r-project.org/web//packages/EGAnet

Golino, H. & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE*, *12*(6), Article e0174035. https://doi.org/10.1371/journal.pone.0174035

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

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*(4), 430–450. https://doi.org/10.1037//1082-989X.6.4.430

Guo, J., Marsh, H. W., Parker, P. D., Dicke, T., Lüdtke, O., & Diallo, T. M. O. (2019). A systematic evaluation and comparison between exploratory structural equation modeling and Bayesian structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(4), 529–556. https://doi.org/10.1080/10705511.2018.1554999

Hakstian, A. R. (1971). A comparative evaluation of several prominent methods of oblique factor transformation. *Psychometrika*, *36*, 175-193. https://doi.org/10.1007/BF02291397

Hakstian, A. R., & Abell, R. A. (1974). A further comparison of oblique factor transformation methods. *Psychometrika*, *39*, 429-444. https://doi.org/10.1007/BF02291667

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, *14*(3), 332–346. https://doi.org/10.1177/1088868310361240

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, P.-H. (2020). lslx: Semi-Confirmatory Structural Equation Modeling via Penalized Likelihood. *Journal of Statistical Software*, *93*(7), 1–37. https://doi.org/10.18637/jss.v093.i07

Huang, P.-H., Chen, H., & Weng, L.-J. (2017). A penalized likelihood method for structural equation modeling. *Psychometrika*, *82*(2), 329–354. https://doi.org/10.1007/s11336-017-9566-9

Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, *26*(3), 395–400. https://doi.org/10.7334/psicothema2013.349

Jacobucci, R., Grimm, K. J., & McArdle, J. J. (2016). Regularized structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(4), 555–566. https://doi.org/10.1080/10705511.2016.1154793

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *Big Five Inventory (BFI)* [Database record]. APA PsycTests. https://doi.org/10.1037/t07550-000

Jöreskog, K. G., & Sörbom, D. (1988). *LISREL 7. A guide to the program and applications* (2nd ed.). International Education Services.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*, 187–200. https://doi.org/10.1007/BF02289233

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptative personality trait model and inventory for DSM-5. *Psychological Medicine*, *42*(9), 1879–1890. https://doi.org/10.1017/S0033291711002674

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2013). *The Personality Inventory for DSM-5 – Adult: Technical Manual*. American Psychiatric Association.

Lai, K. (2019). Creating misspecified models in moment structure analysis. *Psychometrika*, *84*(3), 781–801. https://doi.org/10.1007/s11336-018-09655-0

Liang, X., Yang, Y., & Cao, C. (2020). The performance of ESEM and BSEM in structural equation models with ordinal indicators. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(6), 874–887. https://doi.org/10.1080/10705511.2020.1716770

Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, *34*(3), 347–365. https://doi.org/10.1207/S15327906MBR3403_3

Lorenzo-Seva, U., & Ferrando, P. J. (2020). Unrestricted factor analysis of multidimensional test items based on an objectively refined target matrix. *Behavior Research Methods*, *52*, 116–130. https://doi.org/10.3758/s13428-019-01209-1

Lui, P. P., Samuel, D. B., Rollock, D., Leong, F. T., & Chang, E. C. (2020). Measurement invariance of the five factor model of personality: Facet-level analyses among Euro and Asian Americans. *Assessment*, *27*(5), 887–902. https://doi.org/10.1177/1073191119873978

MacCallum, R. C., Edwards, M. C., & Cai, L. (2012). Hopes and cautions in implementing Bayesian structural equation modeling. *Psychological Methods*, *17*(3), 340–345. https://doi.org/10.1037/a0027131

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

Mair, P., De Leeuw, J., & Groenen, P. J. F. (2021). *smacof: Multidimensional scaling*. *R package version 2.1-2*. https://cran.r-project.org/web/packages/smacof/

Marcoulides, G. A., Drezner, Z., & Schumacker, R. E. (1998). Model specification searches in structural equation modeling using tabu search. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*(4), 365–376. https://doi.org/10.1080/10705519809540112

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2020). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and Set-ESEM:

Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research*, *55*(1), 102–119. https://doi.org/10.1080/00273171.2019.1602503

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially specified target matrices: Applications in exploratory and Bayesian confirmatory factor analysis. *Multivariate Behavioral Research*, *50*(2), 149–161. https://doi.org/10.1080/00273171.2014.973990

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide. Eighth Edition*. Muthén & Muthén.

Nájera, P., Abad, F. J., & Sorrel, M. A. (2022a, June 2). *Is EFA always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory-exploratory continuum*. OSF. https://doi.org/10.17605/OSF.IO/VW4XN

Nájera, P., Abad, F. J., & Sorrel, M. A. (2022b). *wrapFA: A Wrapper for Factor Analysis using lavaan and MplusAutomation*. *R package version 0.0.1*. https://github.com/Pablo-Najera/wrapFA

Nguyen, H. V., & Waller, N. G. (2022). Local minima and factor rotations in exploratory factor analysis. *Psychological Methods*. https://doi.org/10.1037/met0000467

Oltmanns, J. R., & Widiger, T. A. (2020). The five-factor personality inventory for ICD-11: A facet-level assessment of the ICD-11 trait model. *Psychological Assessment*, *32*(1), 60–71. http://dx.doi.org/10.1037/pas0000763

Perry, J. L., Clough, P. J., Crust, L., Earle, K. & Nicholls, A. R. (2013). Factorial validity of the Mental Toughness Questionnaire-48. *Personality and Individual Differences*, *54*, 587–592. https://doi.org/10.1016/j.paid.2012.11.020

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science*, *19*(1), 12–21. https://doi.org/10.1080/1091367X.2014.952370.

R Core Team (2021). *R (Version 4.0) [Computer Software]*. Vienna, Austria: R Foundation for Statistical Computing.

Rindskopf, D. (2012). Next steps in Bayesian structural equation models: Comments on, variations of, and extensions of Muthén & Asparouhov (2012). *Psychological Methods*, *17*(3), 336–339. https://doi.org/10.1037/a0027130

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Roskam, I., Galdiolo, S., Hansenne, M., Massoudi, K., Rossier, J., Gicquel, L., & Rolland, J.-P. (2015). The psychometric properties of the French version of the Personality Inventory for DSM-5. *PLoS ONE*, *10*(7), Article e0133413. https://doi.org/10.1371/journal.pone.0133413

Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. In C. C. Clogg (Ed.), *Sociological Methodology* (pp. 105–129). Jossey-Bass. https://doi.org/10.2307/271030

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling*, *16*(4), 561–582. https://doi.org/10.1080/10705510903203433

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

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, *54*(1), 131–151. https://doi.org/10.1007/BF02294453

Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement*, *71*(1), 95–113. https://doi.org/10.1177/0013164410387348

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. https://doi.org/10.1080/00223891.2018.1449116

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Somma, A., Krueger, R. F., Markon, K. E., & Fossati, A. (2019). The replicability of the Personality Inventory for DSM-5 domain scale factor structure in U.S. and non-U.S. samples: A quantitative review of the published literature. *Psychological Assessment*, *31*(7), 861–877. https://doi.org/10.1037/pas0000711

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*, 371–384. https://doi.org/10.1007/BF02294623

Sorrel, M. A., Aluja, A., García, L. F., & Gutiérrez, F. (2022). Psychometric properties of the five-factor personality inventory for ICD-11 (FFiCD) in Spanish community samples. *Psychological Assessment, 34*(3), 281-293. https://doi.org/10.1037/pas0001084

Sorrel, M. A., García, L. F., Aluja, A., Rolland, J. P., Rossier, J., Roskam, I., & Abad, F. J. (2021). Cross-cultural measurement invariance in the Personality Inventory for DSM-5. *Psychiatry Research*, *304*, Article 114134. https://doi.org/10.1016/j.psychres.2021.114134

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Stromeyer, W. R., Miller, J. W., Sriramachandramurthy, R., & DeMartino, R. (2015). The prowess and pitfalls of Bayesian structural equation modeling: Important considerations for management research. *Journal of Management*, *41*(2), 491–520. https://doi.org/10.1177/0149206314551962

Thimm, J. C., Jordan, S., & Bach, B. (2017). Hierarchical structure and cross-cultural measurement invariance of the Norwegian version of the Personality Inventory for DSM-5. *Journal of Personality Assessment*, *99*(2), 204–210. https://doi.org/10.1080/00223891.2016.1223682

Thurstone, L. L. (1935). *The vectors of mind*. University of Chicago Press.

Tóth-Király, I., Bõthe, B., Rigó, A., & Orosz, G. (2017). An illustration of the exploratory structural equation modeling (ESEM) framework on the Passion Scale. *Frontiers in Psychology*, *8*, Article 1968. https://doi.org/10.3389/fpsyg.2017.01968

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10. https://doi.org/10.1007/BF02291170

Watters, C. A., & Bagby, R. M. (2018). A meta-analysis of the five-factor internal structure of the Personality Inventory for DSM-5. *Psychological Assessment*, *30*(9), 1255–1260. https://doi.org/10.1037/pas0000605

Wei, X., Huang, J., Zhang, L., Pan, D., & Pan, J. (2022). Evaluation and comparison of SEM, ESEM, and BSEM in estimating structural equation models with potentially unknown cross-loadings. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(3), 327–338. https://doi.org/10.1080/10705511.2021.2006664

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *The Journal of Experimental Education*, *80*(1), 26–44. https://doi.org/10.1080/00220973.2010.531299

Wiesner, M., & Schanding, G. T. (2013). Exploratory structural equation modeling, bifactor models, and standard confirmatory factor analysis models: Application to the BASC-2 Behavioral and Emotional Screening System Teacher Form. *Journal of School Psychology*, *51*, 751–763. https://doi.org/10.1016/j.jsp.2013.09.001

Xiao, Y., Liu, H., & Hau, K.-T. (2019). A comparison of CFA, ESEM, and BSEM in test structure analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(5), 665–677. https://doi.org/10.1080/10705511.2018.1562928

Ximénez, C. (2006). A Monte Carlo study of recovery of weak factor loadings in confirmatory factor analysis. *Structural Equation Modeling*, *13*(4), 587–614. https://doi.org/10.1207/s15328007sem1304_5

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. State University of New York Press.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Unpublished doctoral dissertation). University of California, Los Angeles, Los Angeles, CA.

Yuan, K.-H., & Liu, F. (2021). Which method is more reliable in performing model modification: Lasso regularization or lagrange multiplier test? *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(1), 69–81. https://doi.org/10.1080/10705511.2020.1768858

Zhang, B., Luo, J., Sun, T., Cao, M., & Drasgow, F. (2021). Small but nontrivial: A comparison of six strategies to handle cross-loadings in bifactor predictive models. *Multivariate Behavioral Research*. Advance online publication. https://doi.org/10.1080/00273171.2021.1957664

Tables

Table 1

Suggested q-vectors as a function of Factor Correlations and Factor Loadings for a Test Measuring 2 factors using φ = 0.90

					R^2_{jm}			PVAF_{jm}		
ϕ_{12}	λ_{j1}	λ_{j2}	s_{j1}	s_{j1}	{1,0}	{0,1}	{1,1}	{1,0}	{0,1}	{1,1}
0	.10	.50	.10	.50	.010	.250	.260	.038	.962	1
0	.10	.70	.10	.70	.010	.490	.500	.020	.980	1
0	.30	.50	.30	.50	.090	.250	.340	.265	.735	1
0	.30	.70	.30	.70	.090	.490	.580	.155	.845	1
.50	.10	.50	.35	.55	.122	.303	.310	.395	.976	1
.50	.10	.70	.45	.75	.202	.562	.570	.355	.987	1
.50	.30	.50	.55	.65	.303	.423	.490	.617	.862	1
.50	.30	.70	.65	.85	.422	.722	.790	.535	.915	1

Note. ϕ_{12} = correlation between factors 1 and 2; λ_{j1} = factor loading of item *j* on factor 1; s_{j1} = correlation between item *j* and factor 1; {1, 0} = q-vector for item *j* where only factor 1 is specified. For each condition, the PVAF that exceeds φ = .90 is shown in grey. φ = .90 is here used just for illustration purposes. The suggested q-vector would be the one corresponding to the highlighted PVAF.

Table 2

Organization of Factor Analytic Techniques

	Initial Specification	Factor Space	Model Modifications [a]
ICM-CFA	Hypothesis Driven	Restricted	None
BSEM	Hypothesis Driven	Unrestricted	None
CFA_{MI}	Hypothesis Driven	Restricted	Make More Complex
CFA_{S}	Hypothesis Driven	Restricted	Make More Complex
BCFA_{p}	Hypothesis Driven	Restricted	Make Simpler
BCFA_{R2}	Hypothesis Driven	Restricted	Make Simpler
EFA_{T}	Hypothesis Driven	Unrestricted	None
RETAM	Hypothesis Driven	Unrestricted	Complete Refinement
ECFA_{p}	Data Driven	Restricted	Make Simpler
ECFA_{R2}	Data Driven	Restricted	Make Simpler
EFA_{G}	Data Driven	Unrestricted	None

[a] Make More Complex = loadings can be freed, but not fixed; Make Simpler = loadings can be fixed, but not freed; Complete Refinement = loadings can be turned either to specified or unspecified elements in the target matrix.

Table 3

Illustration of the Proportion of Successful Parameter Estimation under N = *1,000*, JK = *8*, FC = *0*, *and* ML = *0.5*

K	*CL*	*CLK*	ICM-CFA	CFA_{MI}	CFA_S	EFA_G	$ECFA_p$	$ECFA_{R2}$	EFA_T	RETAM	BSEM	$BCFA_p$	$BCFA_{R2}$
3	0.15	1	*.00*	.91	.92	1.00	.98	.97	1.00	1.00	.99	.99	.93
3	0.15	2	*.00*	.88	.89	1.00	.99	.96	.94	.94	.79	.98	.95
3	0.30	1	*.00*	1.00	1.00	1.00	1.00	1.00	.96	.88	*.18*	.97	1.00
3	0.30	2	*.00*	.97	.98	.93	.96	.96	*.56*	*.33*	*.07*	.73	.94
5	0.15	1	*.00*	.98	.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.96
5	0.15	2	*.00*	.96	.98	1.00	.98	1.00	1.00	1.00	1.00	1.00	1.00
5	0.30	1	*.00*	1.00	1.00	1.00	.98	1.00	.99	.99	*.22*	.97	.98
5	0.30	2	*.00*	1.00	1.00	1.00	.92	1.00	.99	.99	*.41*	1.00	1.00
		$\overline{PSPE}$	.00	**.96**	**.97**	**.99**	**.98**	**.99**	**.93**	.89	.58	**.96**	**.97**
		$PSPE_{min}$	.00	**.88**	**.89**	**.93**	**.92**	**.96**	.56	.33	.07	**.73**	**.93**
		$PSPE_{max}$	.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; $\overline{PSPE}$ = average proportion of successful parameter estimation; $PSPE_{min}$ = minimum proportion of successful parameter estimation; $PSPE_{max}$ = maximum proportion of successful parameter estimation. *PSPE* values lower than 0.70 are shown in italics. $\overline{PSPE}$ and $PSPE_{min}$ values higher than 0.90 and 0.70, respectively, are shown in bold.

Table 4

Marginal Means of the Bias and RMSE of Major Loadings

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								Bias								
ICM-CFA	.014	.014	.015	.014	.014	.019	.009	–.001	.029	.014	.015	.010	.019	.009	.020	.014
BSEM	.015	.016	.016	.016	.015	.023	.009	.005	.027	.017	.015	.009	.023	.010	.022	.016
CFA$_{MI}$	.009	.007	.007	.007	.009	.012	.003	.000	.015	.011	.004	.007	.008	.005	.010	.008
CFA$_{S}$	.009	.007	.005	.006	.008	.012	.002	.000	.014	.012	.002	.007	.007	.005	.009	.007
BCFA$_{p}$	.007	.007	.006	.008	.006	.011	.003	.000	.014	.008	.005	.005	.008	.004	.009	.007
BCFA$_{R2}$	.006	.006	.004	.007	.004	.009	.002	.000	.011	.008	.003	.006	.005	.003	.008	.005
EFA$_{T}$	–.010	–.007	–.006	–.002	–.013	–.005	–.010	.001	–.016	–.012	–.003	–.006	–.010	–.008	–.007	–.008
RETAM	–.012	–.007	–.006	–.003	–.014	–.006	–.010	.001	–.018	–.014	–.003	–.006	–.011	–.009	–.008	–.008
ECFA$_{p}$	–.008	–.007	–.006	–.005	–.009	–.006	–.008	.001	–.015	–.009	–.005	–.005	–.009	–.006	–.008	–.007
ECFA$_{R2}$	–.007	–.007	–.007	–.005	–.009	–.007	–.006	.000	–.015	–.008	–.006	–.002	–.011	–.006	–.008	–.007
EFA$_{G}$	–.020	–.016	–.015	–.013	–.022	–.016	–.018	.000	–.035	–.019	–.015	–.014	–.021	–.017	–.018	–.017
Total	.000	.001	.001	.003	–.001	.004	–.002	.001	.001	.001	.001	.001	.001	–.001	.003	.001
								RMSE								
ICM-CFA	**.071**	**.058**	.054	**.061**	**.061**	**.072**	**.050**	**.041**	.081	**.071**	.051	**.049**	.073	**.058**	.064	**.061**
BSEM	**.065**	**.053**	**.048**	**.053**	**.057**	**.063**	**.046**	**.040**	**.071**	**.067**	**.044**	**.048**	**.062**	**.053**	**.057**	**.055**
CFA$_{MI}$	.074	**.056**	**.050**	**.056**	.065	.075	**.044**	**.041**	.078	.080	**.038**	**.054**	.066	**.055**	.065	**.060**
CFA$_{S}$	.072	**.056**	**.048**	**.055**	.063	.076	**.041**	**.039**	.078	.082	**.036**	**.051**	.066	**.054**	**.063**	**.059**
BCFA$_{p}$	**.062**	**.051**	**.045**	**.052**	**.053**	**.063**	**.042**	**.039**	**.066**	**.067**	**.038**	**.049**	**.056**	**.050**	**.055**	**.052**
BCFA$_{R2}$	**.065**	**.051**	**.044**	**.052**	**.054**	**.064**	**.042**	**.039**	**.068**	**.070**	**.037**	**.051**	**.055**	**.051**	**.054**	**.053**
EFA$_{T}$	.076	**.057**	**.050**	**.056**	.067	**.072**	.051	**.040**	.082	.081	**.042**	.058	**.064**	.062	**.060**	**.061**
RETAM	.079	**.059**	**.051**	**.058**	.068	.075	.052	**.041**	.086	.085	**.042**	.058	.068	.063	**.063**	.063
ECFA$_{p}$	.072	**.054**	**.048**	**.058**	**.058**	**.071**	**.046**	**.040**	.078	.078	**.040**	**.053**	**.063**	**.056**	**.060**	**.058**
ECFA$_{R2}$	.075	**.052**	**.046**	**.056**	**.059**	**.071**	**.045**	**.039**	.077	.078	**.038**	**.054**	**.062**	**.055**	**.060**	**.058**
EFA$_{G}$	.082	.061	**.053**	**.060**	.070	.077	.053	**.040**	.090	.086	**.045**	.061	.069	.065	.065	.065
Total	.072	.055	.049	.056	.061	.071	.047	.040	.078	.077	.041	.053	.064	.057	.061	.059

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean. The lowest RMSE per condition (differences lower than 0.01 are not considered) is shown in bold.

Table 5

Marginal Means of the Bias and RMSE of Zero-Loadings

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								Bias								
ICM-CFA	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
BSEM	–.009	–.012	–.013	–.016	–.006	–.015	–.008	–.013	–.010	–.008	–.015	–.007	–.016	–.006	–.017	–.011
CFA_{MI}	–.002	–.004	–.004	–.004	–.003	–.004	–.003	–.001	–.006	–.004	–.002	–.002	–.005	–.003	–.004	–.003
CFA_{S}	–.002	–.003	–.003	–.004	–.002	–.004	–.002	–.001	–.005	–.004	–.001	–.001	–.004	–.002	–.003	–.003
$BCFA_{p}$	–.003	–.006	–.008	–.008	–.003	–.007	–.005	–.005	–.007	–.005	–.007	–.003	–.009	–.003	–.008	–.006
$BCFA_{R2}$	–.002	–.003	–.003	–.004	–.001	–.003	–.002	–.002	–.003	–.002	–.002	–.002	–.003	–.001	–.004	–.002
EFA_{T}	.001	–.002	–.003	–.008	.006	–.004	.001	–.013	.011	.005	–.007	–.002	–.001	.003	–.005	–.001
RETAM	.000	–.002	–.003	–.008	.005	–.004	.001	–.014	.011	.004	–.007	–.002	–.002	.002	–.006	–.002
$ECFA_{p}$	.006	.008	.008	.009	.005	.008	.006	.001	.014	.009	.005	.004	.011	.007	.008	.007
$ECFA_{R2}$	.007	.007	.009	.009	.006	.009	.006	.001	.014	.009	.006	.003	.012	.007	.009	.008
EFA_{G}	.017	.016	.016	.019	.013	.016	.017	.003	.030	.020	.013	.010	.022	.016	.017	.016
Total	.001	.000	.000	–.001	.002	–.001	.001	–.004	.005	.002	–.002	.000	.001	.002	–.001	.000
								RMSE								
ICM-CFA	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
BSEM	.035	.039	.041	.040	.037	**.042**	.035	.036	**.041**	**.033**	.043	.032	.045	.036	.041	.038
CFA_{MI}	**.020**	**.028**	**.033**	**.025**	**.028**	**.033**	**.020**	**.012**	**.040**	**.034**	**.019**	**.016**	**.038**	**.023**	**.031**	**.027**
CFA_{S}	**.019**	**.027**	**.030**	**.024**	**.027**	**.033**	**.017**	**.011**	**.039**	**.034**	**.017**	**.015**	**.036**	**.022**	**.029**	**.025**
$BCFA_{p}$	**.022**	**.033**	**.038**	**.032**	**.031**	**.034**	.029	.025	**.038**	**.032**	.031	**.023**	**.040**	**.029**	**.034**	**.032**
$BCFA_{R2}$	**.026**	**.027**	**.028**	**.025**	**.029**	**.030**	**.023**	**.019**	**.034**	**.033**	**.021**	**.023**	**.030**	**.025**	**.028**	**.027**
EFA_{T}	.074	.060	.055	.065	.062	.070	.057	.053	.074	.074	.053	.058	.068	.063	.064	.063
RETAM	.076	.061	.056	.067	.062	.072	.058	.054	.076	.076	.053	.059	.070	.063	.066	.065
$ECFA_{p}$	.044	.042	.043	.044	.042	.050	.037	.030	.057	.052	.035	.036	.050	.043	.043	.043
$ECFA_{R2}$	.041	**.034**	**.035**	.038	**.035**	.045	.028	.023	.050	.047	.027	.029	**.044**	.034	.039	.037
EFA_{G}	.075	.059	.053	.066	.058	.068	.057	.048	.076	.076	.049	.059	.066	.062	.062	.062
Total	.039	.037	.037	.039	.037	.043	.033	.029	.048	.045	.032	.032	.044	.036	.040	.038

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean. The lowest RMSE per condition (ICM-CFA and differences lower than 0.01 are not considered) is shown in bold.

Table 6

Marginal Means of the Bias and RMSE of Cross-Loadings

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								Bias								
ICM-CFA	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.225*	*–.150*	*–.300*	*–.225*	*–.225*	*–.225*
BSEM	*–.119*	*–.089*	*–.072*	*–.096*	*–.088*	*–.109*	*–.076*	*–.095*	*–.090*	*–.125*	*–.063*	*–.068*	*–.117*	*–.096*	*–.089*	*–.093*
CFA$_{MI}$	*–.072*	–.028	–.009	–.038	–.035	*–.052*	–.018	–.033	–.039	*–.057*	–.013	*–.067*	–.007	–.032	–.041	–.036
CFA$_{S}$	*–.070*	–.025	–.010	–.036	–.034	*–.051*	–.019	–.031	–.039	*–.056*	–.014	*–.062*	–.007	–.031	–.039	–.035
BCFA$_{p}$	*–.068*	–.030	–.014	–.041	–.031	*–.058*	–.013	–.032	–.040	*–.061*	–.013	–.046	–.026	–.036	–.036	–.036
BCFA$_{R2}$	*–.068*	–.036	–.021	–.045	–.036	*–.059*	–.022	–.026	*–.056*	*–.065*	–.018	*–.054*	–.028	–.043	–.039	–.041
EFA$_{T}$	–.015	–.017	–.017	–.016	–.017	–.032	–.002	–.034	.002	–.015	–.018	–.009	–.024	–.023	–.010	–.017
RETAM	–.016	–.017	–.017	–.016	–.017	–.033	.000	–.034	.001	–.015	–.018	–.009	–.024	–.022	–.011	–.017
ECFA$_{p}$	–.022	–.003	.003	–.010	–.004	–.018	.003	–.014	.000	–.017	.002	–.017	.003	–.008	–.007	–.007
ECFA$_{R2}$	–.031	–.011	–.002	–.014	–.015	–.020	–.009	–.015	–.014	–.025	–.004	–.032	.003	–.018	–.011	–.014
EFA$_{G}$	.009	.010	.011	.011	.009	.000	.020	–.008	.029	.009	.011	.007	.014	.008	.012	.010
Total	–.063	–.043	–.034	–.048	–.045	–.060	–.033	–.050	–.044	–.060	–.034	–.046	–.047	–.048	–.045	–.047
								RMSE								
ICM-CFA	.225	.225	.225	.225	.225	.225	.225	.225	.225	.225	.225	.150	.300	.225	.225	.225
BSEM	.123	.093	.077	.100	.093	.112	.082	.098	.096	.128	.068	.072	.122	.100	.093	.097
CFA$_{MI}$	.133	.091	.072	.096	.102	.113	.083	.076	.120	.127	.067	.108	.091	.097	.101	.099
CFA$_{S}$	.130	.091	.067	.093	.099	.115	.078	.073	.119	.129	.063	.103	.089	.095	.097	.096
BCFA$_{p}$	.114	.073	.056	.082	.077	.098	.061	.066	.094	.107	.055	.085	.075	.080	.079	.080
BCFA$_{R2}$	.116	.079	.061	.087	.081	.101	.068	.066	.104	.112	.059	.093	.076	.085	.083	.084
EFA$_{T}$	**.062**	**.049**	**.044**	**.049**	**.054**	**.059**	**.044**	**.051**	**.052**	**.061**	**.042**	**.048**	**.055**	**.054**	**.049**	**.051**
RETAM	**.071**	**.054**	**.048**	**.055**	**.060**	**.067**	**.049**	**.052**	.063	.073	**.043**	**.049**	.065	**.057**	**.057**	**.057**
ECFA$_{p}$	.098	.066	.056	.077	.070	.089	.058	.059	.089	.097	**.051**	.075	.072	.071	.075	.073
ECFA$_{R2}$	.107	.070	.056	.079	.076	.093	.062	.061	.095	.105	.052	.086	.069	.076	.079	.078
EFA$_{G}$	.082	.060	**.052**	.066	**.063**	.074	.055	**.048**	.082	.083	**.046**	.061	.068	.064	.065	.064
Total	.115	.087	.074	.092	.092	.105	.079	.080	.104	.114	.070	.085	.099	.092	.092	.092

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean. Absolute biases higher than 0.05 are shown in italics. The lowest RMSE per condition (differences lower than 0.01 are not considered) is shown in bold.

Table 7

Marginal Means of the Bias and RMSE of Factor Correlations

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								Bias								
ICM-CFA	*.109*	*.110*	*.110*	*.132*	*.087*	*.138*	*.081*	*.095*	*.124*	*.124*	*.095*	*.066*	*.152*	*.078*	*.141*	*.109*
BSEM	*.082*	*.078*	*.074*	*.099*	*.054*	*.099*	*.056*	*.076*	*.080*	*.089*	*.067*	*.051*	*.105*	*.053*	*.102*	*.078*
CFA_{MI}	.044	.031	.026	.039	.028	.042	.025	.019	.047	.045	.020	.038	.030	.025	.042	.034
CFA_S	.041	.028	.019	.035	.024	.041	.018	.017	.042	.045	.014	.035	.025	.023	.037	.030
$BCFA_p$	.050	.040	.037	*.054*	.028	*.056*	.027	.031	*.053*	*.055*	.030	.032	*.052*	.028	*.056*	.042
$BCFA_{R2}$	.043	.029	.023	.042	.020	.044	.018	.019	.044	.046	.018	.031	.031	.020	.042	.031
EFA_T	.001	.014	.019	.038	–.016	.024	–.001	*.057*	–.035	–.010	.032	.002	.020	–.002	.025	.011
RETAM	.004	.015	.019	.039	–.015	.026	.000	*.059*	–.035	–.007	.032	.003	.022	.000	.026	.013
$ECFA_p$	–.007	–.019	–.024	–.018	–.016	–.014	–.019	.007	–.043	–.023	–.011	–.008	–.026	–.020	–.014	–.017
$ECFA_{R2}$	–.012	–.020	–.029	–.019	–.022	–.024	–.017	.004	–.046	–.028	–.013	–.001	–.040	–.021	–.020	–.020
EFA_G	*–.061*	*–.055*	*–.052*	*–.057*	*–.054*	*–.052*	*–.059*	.001	*–.115*	*–.080*	–.032	–.046	*–.066*	*–.059*	*–.053*	*–.056*
Total	.026	.023	.020	.035	.011	.035	.011	.035	.011	.023	.023	.018	.028	.011	.035	.023
								RMSE								
ICM-CFA	.136	.125	.122	.143	.113	.158	.097	.122	.134	.146	.110	.087	.168	.101	.155	.128
BSEM	.112	.096	.089	.111	.084	.121	.075	.104	.093	.113	.085	**.075**	.122	.079	.118	.098
CFA_{MI}	.103	**.077**	**.067**	.084	.082	.102	**.061**	.075	.090	.104	**.058**	**.078**	**.087**	**.073**	.092	**.083**
CFA_S	.101	**.075**	**.064**	**.081**	**.079**	.102	**.058**	.072	.088	.104	**.056**	**.075**	**.085**	**.071**	**.089**	**.080**
$BCFA_p$	**.098**	**.074**	**.067**	.086	**.072**	**.098**	**.060**	.076	**.082**	**.098**	**.062**	**.071**	.088	**.070**	**.088**	**.079**
$BCFA_{R2}$	**.097**	**.071**	**.059**	**.080**	**.070**	**.093**	**.056**	.072	**.079**	**.097**	**.055**	**.072**	**.078**	**.069**	**.081**	**.075**
EFA_T	**.090**	**.074**	**.068**	**.074**	.081	**.089**	**.066**	.081	**.073**	**.093**	**.063**	**.071**	**.084**	**.072**	**.082**	**.077**
RETAM	**.094**	**.077**	.070	**.078**	.082	**.092**	.069	.083	**.077**	**.099**	**.063**	**.072**	**.089**	**.074**	**.086**	**.080**
$ECFA_p$	**.093**	**.079**	.075	.088	**.076**	.100	**.065**	**.070**	.095	.105	**.060**	**.074**	.091	.081	**.084**	**.082**
$ECFA_{R2}$	.102	**.079**	.075	.090	.080	.106	**.065**	**.070**	.102	.112	**.059**	**.075**	.095	.081	**.090**	.085
EFA_G	.115	.093	.084	.100	.093	.105	.089	**.061**	.134	.129	.066	.091	.102	.098	.096	.097
Total	.104	.084	.076	.092	.083	.106	.069	.081	.095	.109	.067	.076	.099	.079	.097	.088

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean. Absolute biases higher than 0.05 is shown in italics. The lowest RMSE per condition (ICM-CFA and differences lower than 0.01 are not considered) is also shown in bold.

Table 8

Marginal Means of CFI, RMSEA, and PPP

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								CFI								
ICM-CFA	.848	.851	.852	.875	.826	.866	.835	.817	.884	.800	.901	.870	.831	.858	.842	.850
CFA$_{MI}$	.888	.899	.903	.921	.872	.924	.866	.876	.916	.850	.950	.884	.909	.893	.901	.897
CFA$_{S}$	.890	.901	.904	.923	.874	.925	.872	.880	.916	.850	.946	.887	.910	.894	.902	.898
BCFA$_{p}$	.898	.906	.908	.924	.881	.927	.880	.886	.923	.854	.950	.894	.914	.899	.909	.904
BCFA$_{R2}$	.902	.904	.905	.923	.881	.928	.878	.886	.922	.855	.948	.894	.912	.899	.908	.903
EFA$_{T}$	.912	.913	.913	.934	.891	.945	.882	.895	.931	.872	.953	.903	.922	.908	.918	.913
RETAM	.913	.913	.913	.934	.892	.945	.882	.896	.931	.872	.953	.904	.923	.908	.918	.913
ECFA$_{p}$	.904	.909	.911	.929	.886	.937	.880	.892	.926	.862	.952	.899	.918	.904	.912	.908
ECFA$_{R2}$	.904	.906	.907	.929	.882	.937	.876	.889	.923	.861	.949	.896	.916	.901	.911	.906
EFA$_{G}$	.912	.913	.913	.934	.891	.945	.882	.895	.931	.872	.953	.903	.923	.908	.918	.913
Total	.897	.901	.903	.923	.877	.928	.873	.881	.920	.855	.945	.893	.907	.897	.904	.900
								RMSEA								
ICM-CFA	.064	.064	.064	.067	.061	.070	.058	.066	.062	.058	.070	.056	.072	.060	.068	.064
CFA$_{MI}$	.052	.049	.048	.050	.049	.049	.050	.050	.049	.050	.049	.051	.048	.049	.050	.050
CFA$_{S}$	.052	.049	.048	.050	.050	.049	.050	.050	.049	.050	.049	.051	.048	.049	.050	.050
BCFA$_{p}$	.051	.049	.048	.049	.049	.049	.049	.050	.049	.050	.048	.050	.048	.049	.049	.049
BCFA$_{R2}$	.051	.049	.049	.050	.049	.050	.050	.050	.050	.050	.049	.050	.049	.049	.050	.050
EFA$_{T}$	.052	.052	.052	.051	.053	.052	.053	.054	.050	.052	.052	.054	.050	.053	.052	.052
RETAM	.052	.052	.052	.051	.053	.052	.053	.054	.050	.052	.052	.054	.050	.053	.052	.052
ECFA$_{p}$	.048	.047	.047	.048	.048	.046	.049	.049	.047	.048	.047	.049	.047	.048	.048	.048
ECFA$_{R2}$	.048	.048	.048	.048	.048	.046	.050	.049	.047	.048	.048	.049	.047	.048	.048	.048
EFA$_{G}$	.052	.052	.052	.051	.053	.052	.053	.054	.050	.052	.052	.054	.050	.053	.052	.052
Total	.052	.051	.051	.052	.051	.052	.051	.053	.050	.051	.052	.052	.051	.051	.052	.051
								PPP								
BSEM	.014	.001	.000	.009	.001	.010	.000	.004	.006	.003	.006	.006	.003	.006	.004	.005

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean.

Table 9

Marginal Means of the Determinacy of Factor Score Estimates

	N			*K*		*JK*		*FC*		*ML*		*CL*		*CLK*		
	300	650	1,000	3	5	4	8	0	0.5	0.5	0.7	0.15	0.30	1	2	*Total*
								Empirical reliability								
ICM-CFA	.794	.790	.789	.789	.792	.750	.832	.753	.829	.714	.867	.778	.803	.784	.797	.791
BSEM	.801	.796	.795	.793	.801	.757	.838	.762	.833	.715	.873	.781	.813	.789	.805	.797
CFA_{MI}	.790	.785	.783	.783	.788	.750	.826	.746	.822	.717	.865	.769	.802	.780	.792	.786
CFA_S	.795	.791	.790	.789	.795	.751	.834	.761	.824	.717	.867	.779	.805	.785	.799	.792
$BCFA_p$	.799	.792	.791	.791	.797	.752	.836	.761	.828	.713	.869	.780	.808	.786	.801	.794
$BCFA_{R2}$	.799	.791	.788	.790	.796	.752	.835	.760	.827	.712	.867	.780	.806	.786	.800	.793
EFA_T	.812	.803	.797	.796	.812	.773	.834	.770	.839	.735	.871	.788	.820	.802	.806	.804
RETAM	.813	.803	.797	.796	.813	.774	.834	.770	.839	.735	.871	.788	.820	.803	.806	.804
$ECFA_p$	.799	.789	.783	.787	.795	.751	.829	.763	.821	.712	.865	.778	.804	.785	.796	.791
$ECFA_{R2}$	.800	.786	.783	.785	.794	.749	.829	.762	.819	.712	.864	.778	.801	.785	.794	.790
EFA_G	.813	.800	.789	.784	.817	.773	.827	.767	.835	.734	.865	.785	.816	.799	.802	.801
Total	.801	.793	.789	.789	.800	.575	.832	.762	.829	.720	.868	.780	.809	.790	.800	.795
								True reliability								
ICM-CFA	.744	**.754**	**.756**	**.751**	**.752**	.695	**.808**	.725	.777	**.659**	**.844**	**.753**	.750	**.751**	.752	**.751**
BSEM	**.754**	**.758**	**.759**	**.758**	**.757**	**.707**	**.810**	**.728**	**.789**	**.664**	**.845**	**.752**	**.763**	**.753**	**.762**	**.757**
CFA_{MI}	.742	**.753**	.755	**.751**	.749	**.705**	.801	.720	.778	**.665**	**.848**	.738	**.761**	**.748**	.752	.750
CFA_S	**.749**	**.762**	**.765**	**.760**	**.758**	**.706**	**.812**	**.736**	**.781**	**.665**	**.852**	**.751**	**.767**	**.755**	**.762**	**.759**
$BCFA_p$	**.755**	**.760**	**.762**	**.758**	**.760**	**.708**	**.811**	**.732**	**.788**	**.662**	**.849**	**.752**	**.767**	**.755**	**.763**	**.759**
$BCFA_{R2}$	**.754**	**.761**	**.764**	**.759**	**.761**	**.708**	**.813**	**.734**	**.788**	**.661**	**.851**	**.751**	**.769**	**.755**	**.764**	**.760**
EFA_T	.723	.746	.752	.749	.732	.691	.789	.719	.763	.636	.842	.733	.748	.735	.746	.741
RETAM	.723	.745	.751	.747	.732	.689	.788	.719	.760	.633	.842	.733	.746	.735	.744	.739
$ECFA_p$	.745	**.754**	.754	**.751**	.751	.697	.803	**.731**	.773	.650	**.847**	**.746**	.756	**.747**	.754	**.751**
$ECFA_{R2}$	.740	**.757**	**.759**	**.752**	**.753**	.696	**.807**	**.734**	.772	.650	**.851**	**.747**	.758	**.750**	**.755**	**.752**
EFA_G	.715	.741	.748	.741	.729	.683	.785	.720	.750	.625	.841	.728	.742	.730	.740	.735
Total	.740	.754	.757	.752	.748	.699	.802	.727	.774	.652	.846	.744	.757	.747	.754	.750

Note. *N* = sample size; *K* = number of factors; *JK* = number of items per factor; *FC* = factor correlations; *ML* = magnitude of major loadings; *CL* = magnitude of cross-loadings; *CLK* = number of cross-loadings per factor; *Total* = marginal mean. The highest true reliability per condition (differences lower than 0.01 are not considered) is shown in bold.

Table 10

Results for the Real Data Illustration

	ML	*CL*	*FC*	$\hat{R}^2$	CFI	RMSEA	BIC	*CC*	Somma	ICM-CFA	CFA_{MI}	CFA_S	EFA_T	RETAM	$ECFA_p$	$ECFA_{R2}$
ICM-CFA	.621	.000	.621	.871	.689	.112	88799.8	.999	.201 (.507)							
CFA_{MI}	.633	.144	.364	.854	.880	.075	83104.2	.883	.158 (.296)	.220 (.339)						
CFA_S	.566	.068	.322	.834	.835	.092	84479.5	.893	.151 (.258)	.178 (.391)	.155 (.223)					
EFA_T	.559	.141	.272	.842	.887	.080	82861.5	.977	**.088** (.173)	.197 (.401)	.137 (.206)	.127 (.127)				
RETAM	.555	.143	.318	.844	.887	.080	82861.5	.942	.109 (.202)	.208 (.335)	.137 (.185)	.122 (.154)	**.061** (**.092**)			
$ECFA_p$	.487	.127	.249	.838	.889	.073	82806.6	.933	.181 (.221)	.237 (.454)	.207 (.274)	.154 (.125)	.164 (.118)	.151 (.166)		
$ECFA_{R2}$	.482	.133	.239	.834	.891	.073	82759.7	.958	.179 (.217)	.238 (.475)	.209 (.269)	.159 (.132)	.162 (.119)	.149 (.177)	**.034** (**.043**)	
EFA_G	.484	.146	.214	.833	.887	.080	82861.5	.964	.177 (.205)	.236 (.486)	.207 (.292)	.156 (.166)	.160 (.135)	.147 (.187)	**.040** (**.054**)	**.032** (**.050**)

Note. *ML* = average of major loadings; *CL* = average of cross-loadings; *FC* = average of factor correlations; $\hat{R}^2$ = average determinacy of factor score estimates; Somma = root-mean-square deviation (RMSD) of factor loadings (and factor correlations) with respect to the results from Somma et al. (2019); *CC* = average congruent coefficient across domains and 100 replications. The columns with factor analytic technique names display the RMSD of factor loadings (and factor correlations) between pairs of techniques. RMSD lower 0.100 are shown in bold.

Table 11

Factor loading and correlation matrices of EFA with Geomin and target rotation

	EFA with Geomin rotation (EFA$_G$)					EFA with target rotation (EFA$_T$)				
	Factor loadings									
PID-5 Facets	Neg.	Det.	Ant.	Dis.	Psy.	Neg.	Det.	Ant.	Dis.	Psy.
Anxiousness	**.83**	–.18	–.05	–.02	.03	**.73**	.22	–.03	–.08	.03
Emotional Lability	**.55**	–.36	–.06	.03	.35	**.65**	–.06	–.06	.12	.26
Hostility	.32	–.03	.42	.00	.11	.27	.13	.41	.08	.07
Perseveration	.58	–.01	.07	–.11	.31	.45	.24	.12	–.06	.31
Restricted Affectivity	.02	.61	.23	–.05	.03	–.36	.58	.25	–.09	.09
Separation Insecurity	**.61**	–.39	.16	.03	–.03	**.69**	–.07	.16	.03	–.08
Submissiveness	.45	–.03	.16	.00	–.14	.35	.18	.16	–.06	–.13
Anhedonia	.59	**.45**	.02	.25	–.05	.19	**.80**	–.05	.11	–.10
Depressivity	.67	.19	–.03	.24	.13	.41	.59	–.09	.17	.04
Intimacy Avoidance	.12	**.47**	.01	.03	.12	–.19	**.53**	.01	–.01	.13
Suspiciousness	.48	.11	.11	–.02	.16	.30	.34	.13	–.02	.15
Withdrawal	.37	**.62**	–.01	–.02	.04	–.09	**.77**	.02	–.13	.10
Attention Seeking	.11	–.32	.55	–.09	.10	.27	–.28	.56	.06	.06
Callousness	–.02	.30	.60	.05	.05	–.19	.30	.57	.12	.02
Deceitfulness	.06	.02	**.90**	.05	–.12	.03	.06	**.86**	.15	–.17
Grandiosity	–.04	.12	**.49**	–.29	.12	–.11	.01	**.57**	–.17	.20
Manipulativeness	–.03	.00	**.85**	–.14	–.05	–.03	–.06	**.87**	–.01	–.04
Distractibility	.21	.01	.06	**.30**	.45	.19	.22	–.03	**.41**	.25
Impulsivity	–.03	–.23	.27	**.21**	.40	.16	–.15	.19	**.40**	.21
Rigid Perfectionism	.59	.02	.02	–.59	–.02	.39	.12	.23	–.63	.23
Risk Taking	–.38	–.06	.25	.03	.40	–.23	–.22	.22	.24	.30
Irresponsibility	.01	.06	.45	**.34**	.24	.01	.18	.32	**.47**	.04
Eccentricity	.08	.15	.04	.01	**.67**	.00	.21	.04	.20	**.57**
Perceptual Dysregulation	.21	.01	–.02	–.02	**.74**	.19	.12	.00	.18	**.63**
Unusual Beliefs & Experiences	.00	.07	.02	–.23	**.68**	–.02	.01	.10	–.01	**.67**
Average congruent coefficient	.99	.99	.98	.88	.97	.99	.99	.99	.95	.95
	Factor correlations									
PID-5 Domains	Neg.	Det.	Ant.	Dis.	Psy.	Neg.	Det.	Ant.	Dis.	Psy.
Negative Affectivity	1.00					1.00				
Detachment	.17	1.00				.30	1.00			
Antagonism	.24	.21	1.00			.13	.31	1.00		
Disinhibition	.15	–.03	–.03	1.00		.20	.14	.25	1.00	
Psychoticism	.46	.20	.52	.14	1.00	.32	.40	.49	.27	1.00

Note. Neg. = negative affectivity; Det. = detachment; Ant. = antagonism; Dis. = disinhibition; Psy. = psychoticism. The expected domain for each facet (Somma et al., 2019) is highlighted in gray. Additionally, the facets contributing primarily to each domain as stated in the PID-5 user's manual (see Krueger et al., 2013) are shown in bold.

Figures

Figure 1

Two-dimensional Multidimensional Scaling of Factor Analytic Techniques based on their Parameter Estimates

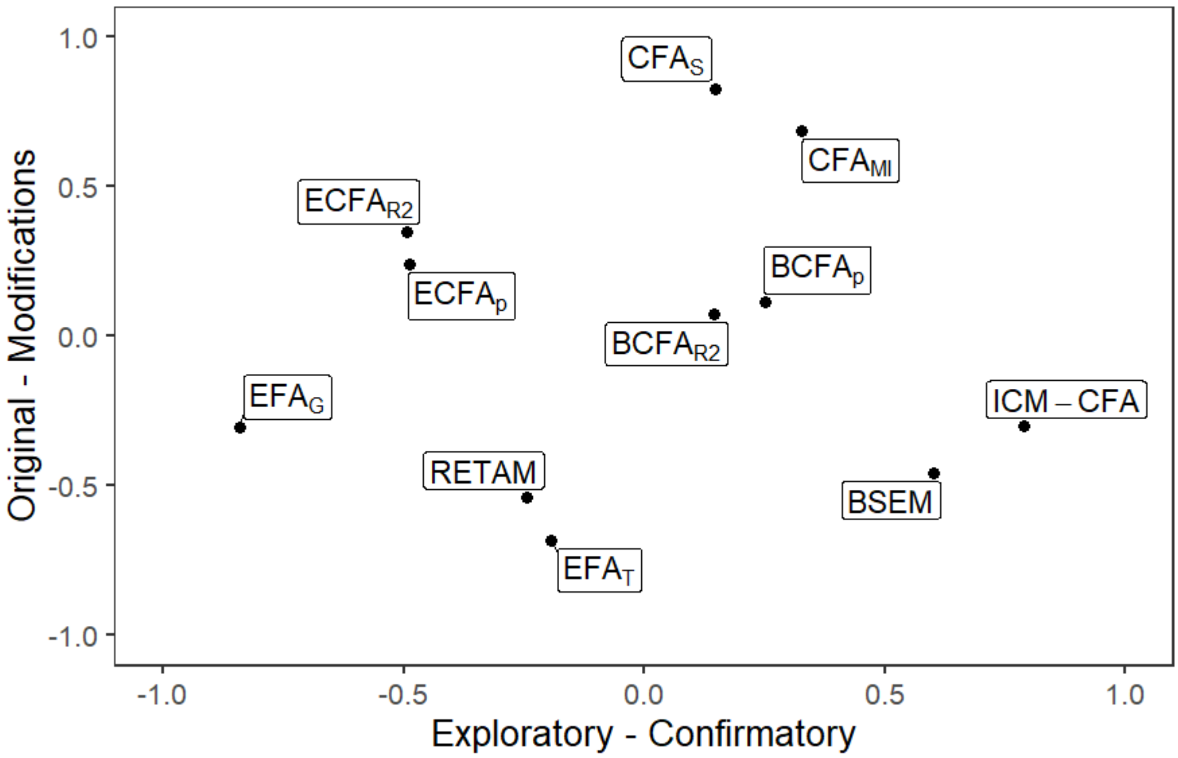
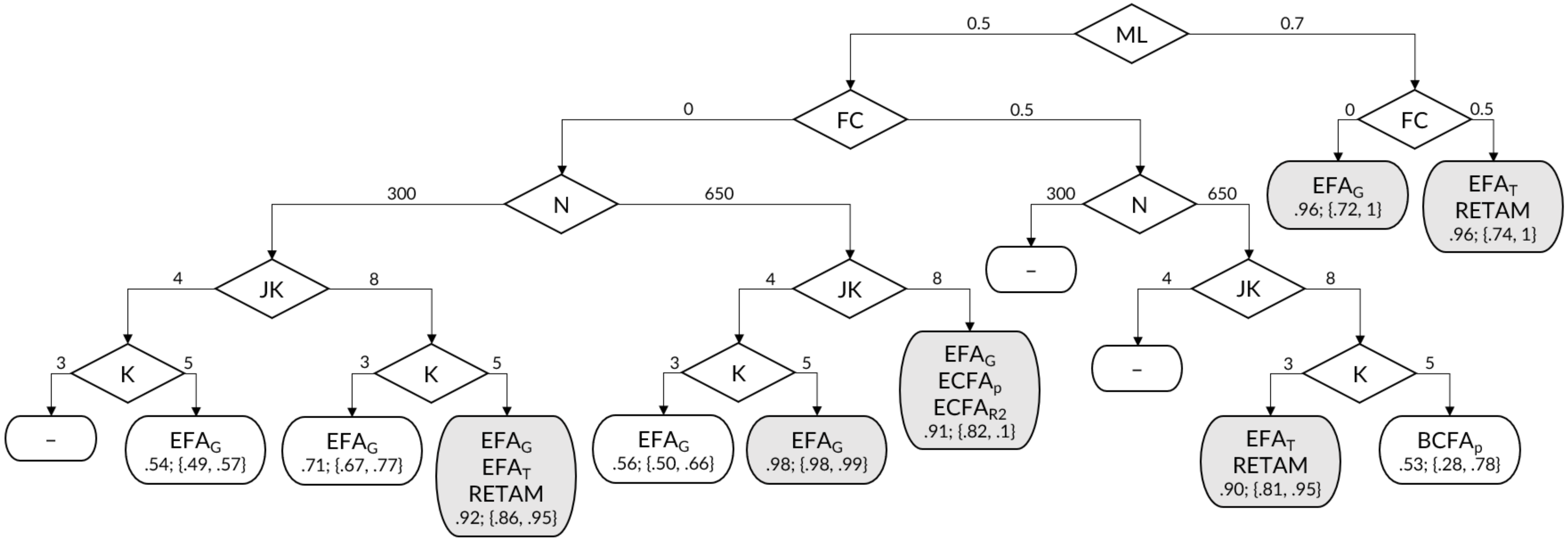


Figure 2

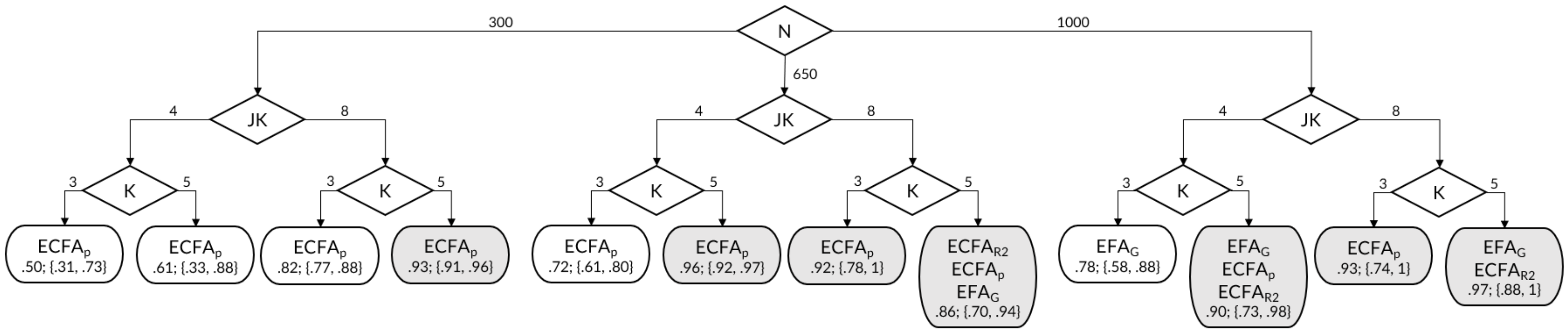
Decision Tree of Factor Analytic Techniques as a Function of Data and Model Conditions based on the Proportion of Successful Parameter Estimation



Note. *ML* = magnitude of major loadings; *FC* = factor correlations; *N* = sample size; *JK* = number of items per factor; *K* = number of factors. Factor analytic techniques that have a $\overline{PSPE} \geq 0.90$ and a $PSPE_{min} \geq 0.70$ are highlighted in gray. Factor analytic techniques that have a $PSPE_{max} \geq 0.50$ are shown in white. A hyphen is shown for those conditions under which no factor analytic technique achieved a $PSPE_{max} \geq 0.50$. Below each factor analytic technique, it is shown the $\overline{PSPE}$; {PSPE_{min}, $PSPE_{max}$}. If more than one factor analytic technique is displayed, the lowest PSPE results among the techniques are shown.

Figure 3

Decision Tree of Data-Driven Factor Analytic Techniques under ML *= 0.7 and* FC *= 0.5*



Note. *ML* = magnitude of major loadings; *FC* = factor correlations; *N* = sample size; *JK* = number of items per factor; *K* = number of factors. Factor analytic techniques that have a $\overline{PSPE} \geq 0.90$ and a $PSPE_{min} \geq 0.70$ are highlighted in gray. Factor analytic techniques that have a $PSPE_{max} \geq 0.50$ are shown in white. Below each factor analytic technique, it is shown the $\overline{PSPE}$; {PSPE_{min}, $PSPE_{max}$}. If more than one factor analytic technique is displayed, the lowest PSPE results among the techniques are shown.