Multivariate Behavioral Research 58.6 (2023): 1072 – 1089

DOI: https://doi.org/10.1080/00273171.2023.2189571

Exploratory Bi-factor Analysis with Multiple General Factors

Marcos Jiménez Henríquez[1,†] Francisco J. Abad[1]

Eduardo Garcia-Garzon[2] Luis Eduardo Garrido[3]

[1]Universidad Autónoma de Madrid (Spain)
[2]Universidad Camilo José Cela (Spain)
[3]Pontificia Universidad Católica Madre y Maestra (Dominican Republic)

ORCID

Marcos Jiménez Henríquez https://orcid.org/0000-0003-4029-6144
Francisco J. Abad http//orcid.org/0000-0001-6728-2709
Eduardo Garcia-Garzon https://orcid.org/0000-0001-5258-232X
Luis Eduardo Garrido https://orcid.org/0000-0001-8932-6063

The published version of this manuscript is in Multivariate Behavioral Research and can be accessed at https://www.tandfonline.com/doi/full/10.1080/00273171.2023.2189571.

† Corresponding author
Departamento de Psicología Social y Metodología
Facultad de Psicología
Universidad Autónoma de Madrid
Ciudad Universitaria de Cantoblanco
28049 Madrid (Spain)
Email: marcosjnezhquez@gmail.com

Declarations

Funding: This research was supported by Grant PSI2017-85022-P (Ministerio de Ciencia, Innovación y Universidades, Spain) and the UAM IIC Chair *Psychometric Models and Applications*.

Conflicts of interest: None.

Material and data availability: The `bifactor` package is available at https://github.com/Marcosjnez/bifactor and the dataset used in this study is available at https://doi.org/10.1371/journal.pone.0133413.

Code availability: All the files necessary to reproduce the simulations can be found at https://osf.io/7aszj.

Abstract

Exploratory bi-factor analysis (EBFA) is a very popular approach to estimate models where specific factors are concomitant to a single, general dimension. However, the models typically encountered in fields like personality, intelligence, and psychopathology involve more than one general factor. To address this circumstance, we developed an algorithm (GSLiD) based on partially specified targets to perform exploratory bi-factor analysis with multiple general factors (EBFA-MGF). In EBFA-MGF, researchers do not need to conduct independent bi-factor analyses anymore because several bi-factor models are estimated simultaneously in an exploratory manner, guarding against biased estimates and model misspecification errors due to unexpected cross-loadings and factor correlations. The results from an exhaustive Monte Carlo simulation manipulating nine variables of interest suggested that GSLiD outperforms the Schmid-Leiman approximation and is robust to challenging conditions involving cross-loadings and pure items of the general factors. Thereby, we supply an R package (`bifactor`) to make EBFA-MGF readily available for substantive research. Finally, we use GSLiD to assess the hierarchical structure of a reduced version of the Personality Inventory for DSM-5 Short Form (PID-5-SF).

Keywords: *Bi-factor analysis, Exploratory factor analysis, Hierarchical structures, Target rotation*

1 Introduction

Bi-factor analysis is an increasingly popular strategy to conceptualize psychological constructs (Reise, 2012). Their distinctive feature is addressing within-item multidimensionality by allowing the indicators to load simultaneously on one orthogonal general factor (e.g., emotional stability) and narrower group factors (e.g., anxiety and depression). In other words, all items share some common variance attributable to a single factor that captures a broader meaning than that of the specific dimensions and is orthogonal to them. It has been argued that this perspective prompts the understanding of complex phenomena like intelligence (Beaujean, 2015), personality (Abad et al., 2018), and psychopathology (Bornovalova et al., 2020), where the data usually display a hierarchical organization, with narrow constructs nested within broader dimensions. As an example, consider the HiTOP model, a new approach to the taxonomy of psychopathology that conceptualizes psychopathological traits across different strata and, ultimately, may conceive a general factor of psychopathology (Kotov et al., 2017). Such hierarchical structures are ubiquitous in psychometric modeling and statistical models like the bi-factor aim to address this important feature.

Currently, the exploratory estimation of bi-factor structures is an active research area with proposals involving the use of analytic rotation criteria (Jennrich & Bentler, 2012, 2011) and target matrices on the factor loadings (Abad et al., 2017; Garcia-Garzon et al., 2019; Lorenzo-Seva & Ferrando, 2019; Waller, 2018). Exploratory bi-factor analysis (EBFA) is a relevant contribution to applied research because real data exhibit complex features (e.g., cross-loadings) that are prone to be misspecified in confirmatory factor analysis (CFA). Usually, CFA is overly restrictive, especially for large factor structures, and such misspecifications severely bias the parameter estimates and undermine model fit indices (Marsh et al., 2014).

Despite these recent advances, a limitation of bi-factor analysis is that it only enables a single general factor, whereas a bi-factor model may include more than one general factor (Giordano et al., 2020) and many instances of psychological assessment involve multiple general factors. As a consequence of this limitation, applied researchers analyzing large factor structures may find themselves constrained to fit an independent bi-factor model to each domain of the data (i.e., analyzing first the items that theoretically load on Neuroticism, then those pertaining to Extraversion, and so on). In this situation, the model misspecifications that EBFA tried to address become a concern again because the items are not allowed to cross-load on general and group factors outside their theoretical domain, with the correlations between the general factors being also ignored. This is highly problematic because, in a domain by domain analysis, item loadings on the theoretical domain would be upwardly biased if they actually load with the same sign on another domain (i.e., interstitial cross-loadings) that is positively correlated with the theoretical one. On the other hand, they would be downwardly biased if the interstitial loadings have opposite signs or the correlation between the domains is negative (Abad et al., 2018). For these reasons, we consider necessary to generalize EBFA to account for multiple general factors (Figure 1), giving raise to exploratory bi-factor analysis with multiple general factors (EBFA-MGF). This generalization accommodates several bi-factor structures within a unique model, presenting a layer of general factors that is orthogonal to the layer of group factors. In EBFA-MGF, all the factor correlations within the same layer of factors and all the cross-loadings would be estimated, offering the opportunity to uncover item complexities and factor correlations that with other methods of analysis would remain hidden, biasing the parameter estimates. In this framework, the group factors bear the same meaning as in the exploratory bi-factor case: they refer to specific content. However, we note an important difference between the traditional bi-factor model and the proposed bi-factor model with multiple general factors. In the former, the general factor is a common dimension affecting all items whereas in the latter, a general factor is conceptualized as a broader dimension that encompasses the indicators pertaining to a subset of group factors. According to this definition, general factors in EBFA-MGF should appear to comprise, at least, two group factors. For instance, in Figure 1 the items $\mathrm{X}_1 - \mathrm{X}_3$ and $\mathrm{X}_4 - \mathrm{X}_6$ are salient indicators of the group factors S_1 and S_2, respectively, and each of these items is also a salient indicator of a broader factor, G_1. In the same manner, the items $\mathrm{X}_7 - \mathrm{X}_9$ and $\mathrm{X}_{10} - \mathrm{X}_{12}$ are salient indicators of the group factors S_3 and S_4, respectively, and each of these items is also a salient indicator of another broader factor, G_2. Thus, there are two general factors defined by the fact that each of them encompasses the salient indicators of two group factors.

Graphically, the bi-factor model with multiple general factors is similar to the two-tier model proposed by Cai (2010). However, the two-tier model assumes a confirmatory simple structure for the group-specific latent dimensions. The model that we propose is also somewhat similar to the two-layer hierarchical model of Tian and Liu (2021), but the latter seeks for simple structure and nested factors within broader factors. On the other hand, EBFA-MGF would estimate a fully exploratory model in which the items loading on the group factors may also load on more than one general factor. Hence, the group factors are not necessarily nested within a single general dimension. For these reasons, we think that the bi-factor model with multiple general factors estimated in EBFA-MGF does not have a clear precedence.

The rest of the manuscript is organized as follows. First, we present the Schmid-Leiman approximation to a bi-factor model with multiple general factors (Schmid & Leiman, 1957). Second, we describe an exploratory approach to estimate the model (i.e., a full-rank bi-factor structure with correlated general factors). Third, we explain the simulation setup and describe the results. Fourth, we illustrate an application of EBFA-MGF in psychopathology using open data. A final discussion of the results, their implications for applied research, and the limitations of the method completes the paper.

1.1 The Schmid-Leiman transformation

The Schmid-Leiman transformation (SL) gives a straightforward approximation to a bi-factor configuration with an arbitrary number of general factors in an exploratory manner (Schmid & Leiman, 1957). It is based on the following hierarchical representation of the empirical correlation matrix $\mathbf{R}$,

$$\mathbf{R} = \mathbf{\Lambda}_1 \mathbf{\Phi}_1 \mathbf{\Lambda}_1^\top + \mathbf{\Psi}_1, \tag{1}$$

$$\mathbf{\Phi}_1 = \mathbf{\Lambda}_2 \mathbf{\Phi}_2 \mathbf{\Lambda}_2^\top + \mathbf{\Psi}_2, \tag{2}$$

where $\mathbf{\Lambda}$, $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote a loading matrix, a correlation matrix among factors, and a diagonal matrix of uniquenesses, respectively. Replacing (2) in (1) and expanding, we have

$$\mathbf{R} = \mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \mathbf{\Phi}_2 \mathbf{\Lambda}_2^\top \mathbf{\Lambda}_1^\top + \mathbf{\Lambda}_1 \mathbf{\Psi}_2 \mathbf{\Lambda}_1^\top + \mathbf{\Psi}_1, \tag{3}$$

which can be arranged as

$$\mathbf{R} = (\mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \mathbf{\Phi}_2^{1/2} \vdots \mathbf{\Lambda}_1 \mathbf{\Psi}_2^{1/2})(\mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \mathbf{\Phi}_2^{1/2} \vdots \mathbf{\Lambda}_1 \mathbf{\Psi}_2^{1/2})^\top + \mathbf{\Psi}_1, \tag{4}$$

where $(\mathbf{X} \vdots \mathbf{Y})$ denotes the column-wise concatenation of matrices $\mathbf{X}$ and $\mathbf{Y}$ with same row dimension. Finally, from (4), we can obtain a bi-factor configuration with multiple (correlated) general factors by setting

$$\mathbf{\Lambda}_{SL} = (\mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \vdots \mathbf{\Lambda}_1 \mathbf{\Psi}_2^{1/2}), \tag{5}$$

$$\mathbf{\Phi}_{SL} = \begin{pmatrix} \mathbf{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{6}$$

The estimation procedure can be summarized in three steps: First, do an exploratory factor analysis (EFA) with the expected number of group factors and apply an oblique rotation to obtain $\hat{\mathbf{\Lambda}}_1$ and $\hat{\mathbf{\Phi}}_1$. Second, do an EFA on $\hat{\mathbf{\Phi}}_1$ by extracting the expected number of general factors and apply an oblique rotation again to get $\hat{\mathbf{\Lambda}}_2$ and $\hat{\mathbf{\Phi}}_2$. In the last step, use the expressions (5) and (6), replacing all the terms by their estimates, to obtain a bi-factor representation of this hierarchical model. As a result, the item loadings on the general factors ($\mathbf{\Lambda}_1\mathbf{\Lambda}_2$) are the sum of their direct effects according to the hierarchical representation, while the item loadings on the group factors ($\mathbf{\Lambda}_1\mathbf{\Psi}_2^{1/2}$) become paths explaining the variance not accounted for by the general factors. Moreover, the group factors are assumed to be orthogonal among them and to the general factors, whereas the correlation between the general factors is estimated ($\hat{\mathbf{\Phi}}_2$).

This transformation may be useful to identify independent cluster structures (McDonald, 2000) and to suggest a configural structure prior to target rotation (Abad et al., 2017; Reise et al., 2011). Unfortunately, SL results in a rank-deficient solution for imposing linear dependencies on the factor loading matrix (Mansolf & Reise, 2016; Waller, 2018). More precisely, the item loadings on the general factors are not independent from the item loadings on the group factors because they share the same ingredients. As a consequence, SL is unable to accurately estimate realistic bi-factor structures including cross-loadings and pure item loadings on the general factor, because the linear dependencies forced by SL are increasingly violated at the population level (Abad et al., 2017; Reise et al., 2011). To our surprise, SL has not been tested in any simulation study contemplating more than one general factor, despite the availability of free software for conducting such analyses (Waller, 2021)[1]. Nevertheless, as we expect the same detrimental performance of SL in the bi-factor case with multiple general factors, we suggest a novel method that aims to perform EBFA-MGF for the first time while efficiently dealing with cross-loadings and pure items. The description of this algorithm, which we have termed the *Generalized Schmid-Leiman iterative Difference-based target rotation* (GSLiD), is given in the next section.

[1]The `SchmidLeiman` function from the `fungible` package (Waller, 2021) already implements the capability of performing this kind of Schmid-Leiman transformation to obtain $\mathbf{\Lambda}_{SL}$ and $\mathbf{\Phi}_2$. They can be accessed via the outputs `$B` and `$Phi2`, respectively.

1.2 The Generalized Schmid-Leiman iterative Difference-based target rotation

We propose an iterative target rotation procedure (GSLiD) that automatically refines the target matrix for the loadings while taking into account the presence of two layers of general and group factors. It can be regarded as a generalization of the SLi and SLiD algorithms developed by Abad et al. (2017) and Garcia-Garzon et al. (2019), which have been applied with success in exploratory bi-factor modeling (Garcia-Garzon et al., 2021), and is devoted to amend the possible misspecification errors in the initial target. This iterative scheme with partially specified targets is not new but was already suggested by Browne (2001, p. 125), and has been recently implemented in other recent algorithms for conducting exploratory factor and bi-factor analyses (Lorenzo-Seva & Ferrando, 2019, 2020; Moore et al., 2015).

Let $\mathbf{A}$ be a $p \times q$ matrix of unrotated factor loadings with p manifest variables and q common factors. The rotation problem is conceptualized as the estimation of a transformation matrix $\mathbf{X}$ such that the rotated factor solution, $\boldsymbol{\Lambda} = \mathbf{A}\mathbf{X}^{-\top}$, minimizes some complexity function to provide a more interpretable loading matrix pattern. When $\mathbf{X}$ is constrained to the oblique manifold of $\mathbb{R}^{q\times q}$ rotation matrices, $\mathcal{OB}(q,q) = \{\mathbf{X} \in \mathbb{R}^{q\times q} : \text{ddiag}(\boldsymbol{\Phi} = \mathbf{X}^\top\mathbf{X}) = \mathbf{I}\}$, where $\text{ddiag}(\mathbf{X})$ returns a diagonal matrix with the diagonal elements of $\mathbf{X}$, the off-diagonal elements of $\boldsymbol{\Phi}$ corresponds to the correlations between the factors.

Until recently, all complexity functions only concerned the rotated loading matrix $\boldsymbol{\Lambda}$. However, Zhang et al. (2019) proposed a new complexity function based on partially specified targets for both factor loadings and factor correlations (i.e., the extended target criterion). This criterion was successfully applied to identify multitrait-multimethod structures where the correlations among trait factors and method factors are freely estimated, but the correlations between them are penalized the more they deviate from zero. The rotation problem posed by the extended target criterion can be defined as finding the solution to

$$\underset{\mathbf{X}\in\mathcal{OB}(q,q)}{\text{argmin}} \quad \frac{1}{2}\|\mathbf{W}_\Lambda \odot (\boldsymbol{\Lambda} - \mathbf{T}_\Lambda)\|^2 + \frac{w}{4}\|\mathbf{W}_\Phi \odot (\boldsymbol{\Phi} - \mathbf{T}_\Phi)\|^2, \tag{7}$$

where $\odot$ is the Hadamard product. $\mathbf{W}_\Lambda$ and $\mathbf{T}_\Lambda$ denote the weight and target matrices for the loading matrix, while $\mathbf{W}_\Phi$ and $\mathbf{T}_\Phi$ bear the analog interpretation for the factor correlations. $\mathbf{T}_\Phi$ must be symmetric and $\mathbf{W}_\Phi$ must be off-diagonal[2] symmetric with nonnegative elements. Lastly, the scalar w represents the relative contribution of the second term in (7) to solve the minimization problem.

In GSLiD, to efficiently rotate a factor solution with an arbitrary number of correlated general factors, we propose to set an initial partially specified target on the factor loadings based on the SL transformation, as described above. Then, the target matrix is updated upon each rotation until it matches the target created in a previous iteration. This update is performed separately for both layers of general and group factors and consists of calculating, for each factor, the mean of the one-lagged differences between the sorted squared normalized loadings. These values are then used as cut-offs to create the new target matrix[3]. In the bi-factor context, such automatic determination of the target has been shown to improve on the demarcation of subjective cut-points in complex structures with many small cross-loadings (Garcia-Garzon et al., 2019). An illustration of this updating method can be found in Table 1 of Garcia-Garzon et al. (2019).

With regard to the targets for the factor correlations, they remain constant in the GSLiD algorithm and must be provided by the researchers according to their theoretical expectations. As an illustrative example, one possibility is to free the correlations among the general factors by fixing their targets to one, fixing to zero the targets for the remainder correlations, and then defining the weight matrix for $\mathbf{\Phi}$, $\mathbf{W}_\Phi$, as the complement of $\mathbf{T}_\Phi$ [4]. These matrices are illustrated in (8) for the case of three general factors and six group factors:

---

[2]The diagonal of $\mathbf{\Phi}$ is a constant vector of ones and therefore is not considered during the minimization.

[3]To encourage the uniqueness of the rotated solution, we additionally checked that the target matrix satisfied the rotational uniqueness conditions in Peeters (2012) in each iteration of the GSLiD algorithm. These conditions ensure that, under oblique rotation, there exists a unique solution when some of the loadings are fixed to zero.

[4]In other published work, it is common to refer to non-specified targets with either asterisks (*) or missing values (NA). Here, such specifications are given by the elements of the weight matrix, where a 0 means the corresponding correlation is freed.

$$\mathbf{T}_\Phi = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{W}_\Phi = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{8}$$

With this setup, we should expect both layers of general and group factors to remain uncorrelated by encouraging the orthogonality of the latter. However, it is important to keep in mind that in any oblique rotation procedure no correlation is guaranteed to be exactly zero. If some noticeable correlations are estimated between the general and group factors after using the extended target criterion, we may simply increase the scalar w to eventually satisfy the orthogonality requirement. This is often desirable for a better interpretation of the model because it allows to disentangle the item variance due to the general and group factors.

The case presented in (8) is an example in which the extended target criterion will be minimized when the group factors are completely orthogonal. Notwithstanding, we would like to remark that the non-orthogonality of the group factors can also be easily accommodated in GSLiD whenever it makes sense from a theoretical point of view. We may just simply change the values of their targets and weights. Another important feature of this rotation criteria is that it does not necessarily encourage a nested structure where the items of a group factor are indicators of a single general factor, as may be done with the Schmid-Leiman procedure. Instead, during the extended target rotation step of the GSLiD algorithm, the items of a group factor may freely load on more than one general factor.

The details of GSLiD and the target updating procedure are outlined in Algorithms 1 and 2, respectively.

Algorithm 1 Exploratory bi-factor analysis with multiple general factors using GSLiD

Inputs: symmetric target matrix for factor correlations, $\mathbf{T}_\Phi$; symmetric weight matrix for factor correlations with non-negative off-diagonal elements, $\mathbf{W}_\Phi$.

Find a Schmid-Leiman solution by calculating the expressions (5) and (6) from a hierarchical factor model.
Set k = 0.
Find $\mathbf{T}_{\Lambda_k}$ and $\mathbf{W}_{\Lambda_k}$ using the loading and factor correlation matrices obtained in the Schmid-Leiman solution as inputs of Algorithm 2.
Set a maximum iteration N.
Estimate an unrotated loading matrix, $\hat{\mathbf{A}}_\mathbf{u}$, by fitting an exploratory factor analysis extracting the total number of common factors (i.e., the sum of general and group factors).
while $k < N$ **do**
    Use $\mathbf{T}_{\Lambda_k}$ and $\mathbf{W}_{\Lambda_k}$ to rotate $\hat{\mathbf{A}}_\mathbf{u}$ by solving the extended target rotation problem (7) and set $\hat{\mathbf{\Lambda}}_{k+1}$ and $\hat{\mathbf{\Phi}}_{k+1}$ as the rotated loading and correlation matrices, respectively.
    Find $\mathbf{T}_{\Lambda_{k+1}}$ and $\mathbf{W}_{\Lambda_{k+1}}$ using $\hat{\mathbf{\Lambda}}_{k+1}$ and $\hat{\mathbf{\Phi}}_{k+1}$ as inputs for Algorithm 2.
    if $\mathbf{T}_{\Lambda_{k+1}}$ is a duplicate (i.e., $\mathbf{T}_{\Lambda_{k+1}} = \mathbf{T}_{\Lambda_j}$ for some $j \leq k$), **then**
        break
    end if
    $k \leftarrow k + 1$
end while

Algorithm 2 Target updating for the loading matrix

Inputs: loading matrix, $\mathbf{\Lambda}$; correlation matrix, $\mathbf{\Phi}$.

1: Create $\mathbf{\Lambda}_{\boldsymbol{g}}$ and $\mathbf{\Phi}_{\boldsymbol{g}}$ by extracting from $\mathbf{\Lambda}$ the columns pertaining to the general factors and, from $\mathbf{\Phi}$, the correlation matrix among them, respectively.
2: Normalize the rows of $\mathbf{\Lambda}_{\boldsymbol{g}}$ by item communalities, $\mathbf{\Lambda}_{\boldsymbol{g}_{\text{norm}}} := \text{ddiag}(\mathbf{\Lambda}_{\boldsymbol{g}}\mathbf{\Phi}_{\boldsymbol{g}}\mathbf{\Lambda}_{\boldsymbol{g}}^\top)^{-1/2}\,\mathbf{\Lambda}_{\boldsymbol{g}}$.
3: Sort the elements of $\mathbf{\Lambda}_{\boldsymbol{g}_{\text{norm}}}^2$ in decreasing order in each column and compute the one-lagged differences by column.
4: Set the mean of each column-vector of one-lagged differences as a column cut-point.
5: Initialize a target matrix $\mathbf{T}_\mathbf{g}$ with same dimensions than $\mathbf{\Lambda}_{\boldsymbol{g}}$.
6: Entries of $\mathbf{T}_\mathbf{g}$ whose corresponding elements in $\mathbf{\Lambda}_{\boldsymbol{g}_{\text{norm}}}^2$ are above the column cut-point are fixed to one and entries below the cut-point are fixed to zero.
7: **if** the identification conditions C1 to C3 defined in Peeters (2012) are not met for $\mathbf{T}_\mathbf{g}$, **then**
8:     the entry corresponding to the smallest non-fixed-to-zero element of each sorted normalized loading column vector is fixed to zero in the target matrix.
9: **end if**
10: Repeat steps 1-9 for the group factors to obtain $\mathbf{\Lambda}_{\boldsymbol{s}}$, $\mathbf{\Phi}_{\boldsymbol{s}}$, and $\mathbf{T}_\mathbf{s}$.
11: Join both target matrices column-wise to obtain the complete target matrix, $\mathbf{T}_\Lambda := (\mathbf{T}_\mathbf{g} \vdots \mathbf{T}_\mathbf{s})$.
12: Define the weight matrix $\mathbf{W}_\Lambda := \mathbf{1}\mathbf{1}^\top - \mathbf{T}_\Lambda$, where $\mathbf{1}$ is a column vector of ones.

2 Methods

To test the estimation accuracy of SL and GSLiD, we ran an extensive simulation involving many variables of interest. The simulation can be considered an extension of the one found in Abad et al. (2017). In this case, two additional variables were considered: (1) the number of general factors and (2) the correlation among the general factors. Thus, nine variables were manipulated in a Monte Carlo simulation to accomplish a fully crossed design that amounts to 7776 conditions, each replicated 50 times. The variables and their levels were: (1) number of general factors (N.GF: 2, 3, 4, 5); (2) correlation between the general factors (COR.GF: 0, 0.5); (3) sample size (N: 500, 1000, 2000); (4) variables per group factor[5] (VAR.GRF: 4, 5, 6); (5) number of group factors defining each general factor (NUM.GRF: 4, 5, 6); (6) cross-loadings among the group factors (CROSS.GRF: no, yes); (7) factor loadings on the group factors (LOAD.GRF: low, medium, high); (8) factor loadings on the general factors (LOAD.GF: low, medium, high); and (9) pure indicators of the general factors (PURE.GF: no, yes).

The factor loadings were generated from .30 to .50 for the low loadings condition, from .40 to .60 for the medium condition, and from .50 to .70 for the high condition. In every case, the loadings ranged by equal increments across the indicators of each group factor (e.g., for the low condition with four items by group factor, the population factor loadings were .30, .37, .43, and .50). When cross-loadings were present, the item with the greatest loading on each group factor had a cross-loading of .40 in another group factor. Moreover, to maintain the communality constant, a small value was subtracted from the remaining non-zero item loadings. In addition, pure indicators in the general factors were determined by decreasing the loading of the middle item of each group factor to .01 (e.g., the second item of each group factor in a four-item condition and the third item in a five-item condition) and increasing the loading on the general factor in order to maintain, again, the initial communality.

[5]Please, note that VAR.GRF indicates the ratio between the total number of items to the total number of group factors and not the number of variables that are indicators of each group factor. The last interpretation would only be correct for structures without pure items.

With this simulation, we tried to investigate the stability of the methods in the presence of two well-known disturbances of the simple structure, namely cross-loadings between the group factors and pure item loadings on the general factors. The combinations of these variables recreate the four types of structures investigated in Abad et al. (2017): (IC) Independent cluster structure: neither cross-loadings nor pure indicators are present; (ICB) Independent cluster basis: cross-loadings but not pure indicators are present; (ICP) Independent cluster pure: pure indicators but not cross-loadings are present; and (ICBP) Independent cluster pure basis: both cross-loadings and pure indicators are present. A simulated pattern for the IC, ICB, ICP, and ICBP conditions is displayed in Table 1.

The performance of the SL and GSLiD methods were compared in two outcomes: the average of the Tucker's factor congruence coefficients (ACC; Burt, 1948) between the simulated and estimated factor loadings and the root mean square error between the true and estimated correlations among the general factors ($\hat{\mathbf{\Phi}}_g$ RMSE),

$$\mathrm{ACC} = \frac{1}{q} \sum_j \frac{\sum_i \hat{\lambda}_{ij} \lambda_{ij}}{\sqrt{\sum_i \hat{\lambda}_{ij}^2 \sum_i \lambda_{ij}^2}}, \quad \hat{\mathbf{\Phi}}_g \text{ RMSE} = \sqrt{\sum_{i>j} \frac{\left(\phi_{g_{ij}} - \hat{\phi}_{g_{ij}}\right)^2}{g(g-1)/2}}, \tag{9}$$

where g denotes the number of general factors.

Congruence coefficients greater than .95 were taken to indicate an adequate level of similarity between factor loadings (Lorenzo-Seva & Berge, 2006) and root mean square errors smaller than .05 were considered good levels of misfit.

For each condition, we generated 50 population structures from which a random sample was drawn from a multivariate normal distribution. ANOVAs estimating up to third-order interactions among all the variables, treated as factors, were carried out for each combination of outcome and method. The partial omega squared (Ω^2_{prtl}) was then used as an effect size measuring the importance of each coefficient. Following the benchmarks proposed by Cohen (1988) for eta squared effect sizes, we differentiated between small ($\Omega^2_{prtl} = .01$), medium

($\Omega^2_{prtl} = .06$) and large ($\Omega^2_{prtl} = .14$) effect sizes.

Unweighted least squares estimation was applied to fit the factor models. When Heywood cases were encountered, minimum rank factor analysis was performed to ensure that positive uniquenesses were estimated. The quartimin criterion was applied to rotate the first and second-order solutions for SL. To attain a global minimum in the rotation step within each target iteration, we generated ten random orthogonal matrices as starting values and selected the solution which produced the smallest objective function. These orthogonal matrices were obtained as the Q factors of the QR decompositions of matrices with random standard normal deviates. The maximum number of target iterations in the GSLiD algorithm was set to 100 to guarantee that the estimated loading matrix converged to an optimal target specification (when it existed). Nevertheless, convergence failure may still occur when the updated target is a duplicate of a previous one that is different from the target computed in the last iteration. In this case, the algorithm would enter an endless loop. When such a situation was identified, we decided to retain the solution obtained in the current iteration. To check whether these solutions were suboptimal compared to the solutions which attained convergence, we ran two analyses of variance, one for each outcome (ACC and $\hat{\mathbf{\Phi}}_g$ RMSE), using the convergence of the GSLiD algorithm as an additional factor to the nine variables listed above.

All simulations were performed in `R` (R Core Team, 2018) under the 4.0.3 version. The models were fitted using the `bifactor` package, version 0.1.0. The congruences between the true and estimated factor loading matrices were calculated by matching both via least squares, using the `faAlign` function from the `fungible` package (Waller, 2021), version 2.2. The ANOVAs were executed with the `aov` function and treating all the variables as factors. A development version of the `bifactor` package can be downloaded from https://github.com/Marcosjnez/bifactor and the necessary files to reproduce the simulations are available at https://osf.io/7aszj.

3 Results

Few Heywood cases were encountered (< 0.08%), and no rotation convergence failure for the extended target criterion was observed. However, 2.5% of the simulations resulted in recurrent target iterations without convergence. Nonetheless, the ANOVAs on both outcomes did not reveal an effect of the convergence of the GSLiD algorithm as a factor ($\Omega^2_{prtl} = .00$ for ACC and $\hat{\mathbf{\Phi}}_g$ RMSE), so we retained all replicates in subsequent analyses.

Table 2 contains the marginal outcomes for each variable level. Marginal ACCs were high for the GSLiD method across all the variables except for some unfavorable conditions such as low loadings on the group factors (ACC [LOAD.GRF = low] = .923) and the minimum sample size condition (ACC [N = 500] = .934). In total, 19 of the 25 levels considered in the simulation resulted in an ACC greater than .95 for GSLiD, contrasting with the four observed for SL. In fact, GSLiD performed better or equal (ACC [PURE = no] = .966) than SL across all the variable levels. Overall, the sample size, the number of items per group factor and the loadings' magnitude on the group and general factors were positively related to the ACC, whereas the number of general and group factors, cross-loadings and pure items diminished the ACC. Conversely, the correlation among the general factors affected the performance of neither method. The results of the ANOVA on the ACC (Table 3) confirmed that GSLiD was substantially less sensitive than SL to most of the variables, with the latter being largely influenced by the presence of pure items and cross-loadings (Ω^2_{prtl} [PURE.GF] = .90; Ω^2_{prtl} [CROSS.GRF] = .80), which were also involved in several high two-way interactions. Whereas SL slightly overcame GSLiD in the independent cluster structure (IC: ACC [SL] = .975; ACC [GSLiD] = .965), it provided worse results in the remaining structures. Figure 2 displays the third-order interaction between pure items, cross-loadings and the number of variables per group factor. GSLiD was stable in all the conditions, except under ICBP structures with four indicators per group factor, while SL underperformed in the presence of pure items (ICP), especially when they occurred simultaneously with cross-loadings in the ICBP structures (Ω^2_{prtl} [CROSS.GRF $\times$ PURE.GF] = .62).

Concerning the recovery of the correlations among the general factors, all marginal $\hat{\mathbf{\Phi}}_g$ RMSEs were much smaller for GSLiD than SL, improving the correlation estimates across all the four structure types. In total, 23 out of 25 marginal RMSEs were smaller than .05 for GLSiD, while SL only produced an average RMSE below this threshold under the orthogonal general factors level ($\hat{\mathbf{\Phi}}_g$ RMSE [COR.GF = 0] = .031). Increasing sample sizes also reduced the $\hat{\mathbf{\Phi}}_g$ RMSE while increasing the number of general factors undermined the accuracy of the correlation estimates. Remarkably, all these effects were stronger for SL. The magnitude of the loadings on the group factors increased the $\hat{\mathbf{\Phi}}_g$ RMSE for SL and did not affect GSLiD. In contrast, the loadings' magnitude on general factors affected GSLiD but not SL. Concretely, the $\hat{\mathbf{\Phi}}_g$ RMSE diminished progressively with higher loadings on the general factors. Finally, the effect of the number of group factors and the number of items per group factor were small. According to the ANOVA, the most important variable affecting the accuracy of the methods was COR.GF ($\Omega^2_{prtl} \geq .55$), indicating that the $\hat{\mathbf{\Phi}}_g$ RMSE was much smaller for both methods when estimating true zero correlations. The presence of cross-loadings affected SL (Ω^2_{prtl}[COR.GF $\times$ CROSS.GRF] = .34) while pure items influenced GSLiD (Ω^2_{prtl}[COR.GF $\times$ PURE.GF] = .28). However, the role of these variables was different in each method, with cross-loadings impairing SL (Figure 3a) and pure items benefiting GSLiD (Figure 4a). Additionally, the interaction COR.GF $\times$ N.GF revealed that SL is sensitive to the number of general factors when they are correlated (Figure 3b). As a downside, the interaction COR.GF $\times$ LOAD.GF exposed that GSLiD was more susceptible to the magnitude of the loadings on correlated general factors (Figure 4b), with smaller magnitudes worsening the estimation.

3.1 Personality Inventory for DSM-5 Short Form

The Personality Inventory for DSM-5 Short Form (PID-5-SF; Maples et al., 2015) is an instrument that aims to measure maladaptive personality features on 25 traits and five domains using 100 items, four by trait. However, the American Psychiatric Association instructs clinicians to measure the five domains using 15 traits, three per domain[6]. The domains (G) and traits (S) are listed as follows: Negative Affect (G1), Emotional Lability (S1), Anxiousness (S2) and Separation Insecurity (S3); Detachment (G2), Withdrawal (S4), Anhedonia (S5) and Intimacy Avoidance (S6); Antagonism (G3), Manipulativeness (S7), Deceitfulness (S8) and Grandiosity (S9); Disinhibition (G4), Irresponsibility (S10), Impulsivity (S11) and Distractibility (S12); Psychoticism (G5), Unusual Beliefs (S13), Eccentricity (S14) and Perceptual Dysregulation (S15).

To investigate this structure, we selected the PID-5-SF items that belong to the factors listed above, retaining a total of 60 items[7]. Data of 2532 participants from the French validation of a larger inventory (Roskam et al., 2015) were employed. To assess the hierarchical organization of their data, Roskam et al. (2015) diagnosed the presence of 5 general factors using Goldberg's Bass-Ackwards method (Goldberg, 2006). However, the Bass-Ackwards is not a truly hierarchical method but a way of summarizing solutions for different number of factors. In contrast, we assessed the hierarchical organization of the data using hierarchical exploratory graph analysis (hierEGA), a method that has shown to be highly accurate in a recent simulation (Jimenez et al., 2022). In the end, hierEGA suggested 16 group factors and 5 general dimensions, concurring in the number of general factors with the Goldberg's Bass-Ackwards method. However, only one item loaded primarily on the additional factor estimated with GSLiD. Therefore, we decided to refit the model using 15 group factors, which is the number expected by theory. The polychoric correlation matrix was used as input and the oblimin criterion was employed to obtain the first and second-order solutions for SL. We freed the correlations between the general factors and fixed to zero the targets for all the remaining correlations. The GSLiD algorithm detected an optimal target after eight

[6]See the 8th page of the APA template, which can be downloaded from https://osf.io/b9rjh/.

[7]The items we retained were 122, 138, 165, 181 (Emotional Lability); 79, 109, 130, 174 (Anxiousness); 50, 127, 149, 175 (Separation Insecurity); 82, 136, 146, 186 (Withdrawal); 23, 26, 124, 157 (Anhedonia); 89, 120, 145, 203 (Intimacy Avoidance); 107, 125, 162, 219 (Manipulativeness); 53, 134, 206, 218 (Deceitfulness); 40, 114, 187, 197 (Grandiosity); 129, 156, 160, 171 (Irresponsibility); 4, 16, 17, 22 (Impulsivity); 118, 132, 144, 199 (Distractibility); 106, 139, 150, 209 (Unusual Beliefs); 25, 70, 152, and, 205 (Eccentricity); 44, 154, 192, 217 (Perceptual Dysregulation).

iterations, each performing ten rotations with random starting orthogonal matrices to avoid local minima. Neither Heywood cases nor rotation convergence failures were encountered. The estimated loading and factor correlation matrices are displayed in Tables 4 and 5 from Appendix B, respectively.

With respect to the hypothesized hierarchical structure of the PID-5-SF, most items presented medium to high loading magnitudes on their expected general and group factors. The target matrix obtained by the GSLiD algorithm agreed with the theoretical pattern of the PID-5-SF 90% of the time when assigning a 1 to a factor loading. 50 items (83.3%) and 54 items (90%) primarily loaded on their expected general and group factors, respectively. The indicators that did not conform to this pattern were items 1, 2, 37, 38, 39, 40, 53, 54, 55 and 56, in the first case, and items 31, 38, 40, 51, 57 and 60, in the second. Four items were pure indicators of a general factor (6.7%; items 14, 15, 28 and 52) and three items cross-loaded on another group factor (0.5%; items 3, 57 and 59). Three Detachment domain (G2) items cross-loaded on Negative Affect (G1) and two from Negative Affect (G1) cross-loaded on Disinhibition (G4). Also, eight items pertaining to two group factors switched the domain on which they were expected to load: items 37, 38, 39 and 40 loaded on Antagonism (G3) instead of on Disinhibition (G4), and items 53, 54, 55 and 56 loaded on Disinhibition (G4) instead of on Psychoticism (G5). This novel result may suggest that Irresponsability (S10) and Eccentricity (S14) could be traits related to different domains than previously thought.

Finally, the correlations between the general factors were moderate, ranging from .12 to .64, while the correlations between the layers and among the group factors remained negligible (i.e., all the estimated correlations were below .10). Thereby, we can conclude that the underlying structure of the PID-5-SF is compatible with a bi-factor structure with multiple general factors and low-to-moderate loadings and factor correlations.

Alternative analyses to GSLiD are also possible upon the availability of a theory supporting a particular factor structure, like the case at hand. For instance, we could perform a plain orthogonal target rotation using the presumed PID-5-SF pattern to build the target matrix.

A shortcoming of this approach is that ignoring the factor correlations between the general factors will result in the estimation of many spurious cross-loadings if they are truly correlated. To avoid this problem, we can simply replace the target criterion with the extended target criterion to encourage the orthogonality of the group factors; or, even better, we may iterate the PID-5-SF target in a similar scheme as GSLiD does[8]. However, all these analyses rest on a theoretical target that does not always exist in practice and that, when available, may provide unstable solutions in structures with low communalities (Myers et al., 2015). Furthermore, rotations involving theoretical targets may produce overconfidence in desired pattern structures that are different from the true ones (Hurley & Cattell, 1962). Moore et al. (2015) also warned that iterating from an initial theoretical target is still at risk of validating a wrong theory and that beginning from an empirically-defined target should be preferred. In these regards, GSLiD offers a solution to preclude such confirmation bias, facilitating the discovery of misspecifications in the theory (i.e., identifying items landing on different group and general factors than expected by the theory)[9].

4 Discussion

Until now, researchers have been restricted to separately analyze general dimensions to build complex models, ignoring the presence of cross-loadings and factor correlations across the structures of different general factors. Consequently, current models may not resemble important aspects of the hierarchical structures commonly encountered in many fields like intelligence, personality, and psychopathology, where narrow constructs are usually nested within broader dimensions. Therefore, we propose EBFA-MGF, an extension of EBFA that estimates factor structures involving multiple general factors. A key feature of EBFA-MGF is

[8] All the code to execute these alternative analyses can be found at https://osf.io/tb2kh/.

[9] At the time of exploring alternative analyses, we noticed that the `SLi` function from the `fungible` package permits the estimation of a bi-factor model with multiple general factors, generalizing the SLi method proposed by Abad et al. (2017) for the bi-factor case. However, at difference with GSLiD, the former does not use the extended target criterion to avoid estimating the factor correlations between the general and group factors nor use the improved cut-off determination developed by Garcia-Garzon et al. (2019).

that it estimates a fully exploratory model, allowing all items to cross-load in both layers of general and group factors. Furthermore, factors within the same layer may be allowed to freely correlate among them. These are important advantages of EBFA-MGF over confirmatory factor analysis, which usually leaves these real data features misspecified resulting in biased parameter estimates and unacceptable model fit indices (Marsh et al., 2014).

We developed an algorithm (GSLiD) to reliably perform, for the first time, this kind of analysis. As emphasized by Marsh et al. (2020), confirmatory factor analysis lacks the flexibility to identify cross-loadings, while exploratory factor analysis may lack parsimony. On the other hand, Zhang et al. (2019) stated that target rotation can be considered a procedure that lies between CFA and EFA and we think this view encourages the utility of GSLiD as a reliable method capable of uncovering complex factor structures involving several general factors in a parsimonious way. The flexibility of GSLiD lies in the fact that the model estimation is completely exploratory (i.e., all parameters are estimated), with the orthogonality between the general and group factors being approximated by fixing to zero the targets related to such factor correlations. Another flexibility of this method lies on the possible patterns of loadings that it can estimate. Whereas the initial target created from the SL solution usually has a nested indicator structure (i.e., the items of a group factor are mainly indicators of a single general factor), the extended target rotation does not penalize non-nested indicator structures but allows complex patterns with the indicators of a group factor loading on different general factors.

Overall, the Monte Carlo results showed GSLiD was less sensitive than SL to all the variables considered in the simulation. Furthermore, GSLiD largely outperforms SL not only by demonstrating a good performance across most conditions but stability under complex structures with cross-loadings and pure items (ICBP). In contrast, SL retrieves good average congruence coefficients in IC and ICB structures but breaks down once pure items are present (ICP), especially when they concur with cross-loadings in ICBP structures. The reason behind the defective performance of SL is that it cannot adequately reproduce full-rank bi-factor patterns where items load on general factors but do not load on the group factors: according to SL, an item loading on a general factor is a linear combination of the item loadings on the factors from the first-factor solution weighted by these factor loadings on such general dimension; thereby, when the latter loadings are small, the former loadings must also be small. Consequently, spurious loadings on the group factors may be estimated to account for the variance explained by the general factors. This problem is exacerbated in the presence of cross-loadings between the group factors because they increase item communalities, and thus favors higher item loadings on the general factors. Hence, SL is also incompatible with a modest item loading on a general factor but several item loadings on the group factors. These shortcomings of SL may induce a biased initial target matrix when trying to estimate an ICBP structure, explaining why GSLiD displayed a modest performance in this structure type under the conditions involving four items per group factor. Regarding the estimation of the general factor correlations, the estimates provided by SL were also increasingly inaccurate in the presence of cross-loadings but not in the presence of pure items. In contrast, cross-loadings bore no effect for GSLiD, but pure items contributed to improve the estimates of the general factor correlations.

Our simulation study has several strengths. On the one hand, it is the first in investigating the performance of exploratory methods in bi-factor situations involving more than one general factor, which is of interest for many fields in individual differences. On the other hand, we manipulated many variables of interest to achieve a comprehensive understanding of the strengths and pitfalls of the methods that were tested. Additionally, the simulations were executed assuming nothing was known about the underlying factor structure beyond the number of group and general factors, which is desirable to avoid model misspecification. In this situation, our results reveal that the determination of the initial target based on a Schmid-Leiman transformation is justified. However, the performance of GSLiD was not investigated under dimensionality misspecification, so we advise caution when GSLiD is used and the number of general factors and group factors are unknown. In a similar simulation study, we devised a method to assess the hierarchical structure of bi-factor data with multiple general factors, termed hierarchical exploratory graph analysis (hierEGA, Jimenez, Abad, Garcia-Garzon, Golino, et al., 2022). This hierarchical version of EGA displayed high accuracies when estimating the number of group factors and yielded a close to perfect hit rate with respect to the number of general factors. Therefore, we recommend to assess the dimensionality of the data with hierEGA[10]. Other variables worth considering for oncoming simulations and not covered here are cross-loadings between general factors, correlations among group factors, and systematic noise in the form of correlated errors. Another limitation is that we did not study the behavior of fit indices nor compare the bi-factor model with multiple general factors to other competing models.

We also remark the unexplored possibility of supplying a custom initial target for the GSLiD algorithm in the case that more information is available about the loading matrix pattern. This possibility is already implemented in the `bifactor` package but the benefits of such custom initial targets are still unknown and should consider the problem of of confirmation bias (Hurley & Cattell, 1962; Moore et al., 2015). In this case, and following the results of our simulation, we would recommend to specify at least four targets per column since four salient indicators per group factor resulted in good factor congruences in most of the investigated conditions. However, this number depends on the complexity and size of the factor structure at hand.

Another possibility worth studying, although uncommon in the psychometric literature, involves estimating an additional general factor in which all items load, resulting in a three-layer bi-factor model akin to the one proposed in Tian and Liu (2021). This new factor would be orthogonal to any other, so that the item variance that it explains can be differentiated from that of the remaining factors. Implementing such a model with GSLiD would be easy, since it only requires to generalize further the Schmid-Leiman transformation to extract a third-level general factor and to adjust the construction of the target matrices to accommodate

[10]The hierEGA method for dimensionality assessment is already available via the function `hierEGA` from the `EGAnet` package (Golino & Christensen, 2022), version 1.1.0.

this new factor. Interestingly, the function `SchmidLeiman` from the `fungible` package already allows this generalization of the Schmid-Leiman approximation but no simulation research has been conducted yet to evaluate its performance.

The robustness of GSLiD can already be exploited to uncover features that remain hidden under hierarchical representations and modeling limitations. The portraying feature of GSLiD consists of automatically updating the target for the factor loading matrix, so that the initial target misspecifications can be empirically resolved to successfully identify cross-loadings and pure items. Therefore, it is well suited to study the large and complex scale structures encountered nowadays in intelligence, personality, and psychopathology research, where bi-factor models with multiple general factors have not been explored enough. For instance, consider the Hierarchical Taxonomy of Psychopathology (Kotov et al., 2017), a new classification system that considers the dimensional nature of psychopathology to increase the reliability of diagnoses. Ultimately, it is proposed as an alternative to the DMS classification scheme by addressing the need to establish clear boundaries between psychopathological conditions. For this aim, the GSLiD algorithm offers a reliable method to identify cross-loadings between items referring to different maladaptive traits and, more broadly, to different spectra. Moreover, its ability to identify pure items may also become useful to distinguish exclusive indicators of spectra. To illustrate how EBFA-MGF can be done with GSLiD in this context, we analyzed a real dataset concerning maladaptive personality traits and compared the estimated multiple bi-factor pattern to the presumed structure of the PID-5-SF. The results showed that, although we found considerable agreement between the theoretical and the estimated factor patterns, there were important cross-loadings and pure items that should not be ignored. Yet, the most interesting result implied that many items loaded on group factors related to a different general dimension than expected by the theory, suggesting that a different configural pattern should also be assessed when analyzing data from the PID-5-SF. Indeed, another important feature of GSLiD is beginning the iterative process from a target that is empirically built. This route guards against the confirmation bias that may happen when specifying a theoretical target. As warned by Moore et al. (2015), it is sometimes easy to incorrectly support the configural structure of a theory when the same theory was used to build the target matrix. In this regard, GSLiD may become specially useful for detecting misspecifications in the theoretical model structure by identifying items loading on different factors than expected in theory.

References

Abad, F. J., Garcia-Garzon, E., Garrido, L. E., & Barrada, J. R. (2017). Iteration of partially specified target matrices: Application to the bi-factor case. *Multivariate Behavioral Research*, *52*(4), 416–429. https://doi.org/10.1080/00273171.2017.1301244

Abad, F. J., Sorrel, M. A., Garcia, L. F., & Aluja, A. (2018). Modeling general, specific, and method variance in personality measures: Results for ZKA-PQ and NEO-PI-R. *Assessment*, *25*(8), 959–977. https://doi.org/10.1177/1073191116667547

Absil, P.-A., Baker, C. G., & Gallivan, K. A. (2007). Trust-region methods on Riemannian manifolds. *Foundations of Computational Mathematics*, *7*(3), 303–330. https://doi.org/10.1007/s10208-005-0179-9

Absil, P.-A., & Gallivan, K. A. (2006). Joint diagonalization on the oblique manifold for independent component analysis. *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, *5*, V–V. https://doi.org/10.1109/ICASSP.2006.1661433

Absil, P.-A., Mahony, R., & Sepulchre, R. (2010). Optimization on manifolds: Methods and applications. In M. Diehl, F. Glineur, E. Jarlebring, & W. Michiels (Eds.), *Recent Advances in Optimization and its Applications in Engineering* (pp. 125–144). Springer. https://doi.org/10.1007/978-3-642-12598-0_12

Absil, P.-A., Mahony, R., & Sepulchre, R. (2009). *Optimization algorithms on matrix manifolds*. Princeton University Press. https://www.degruyter.com/document/doi/10.1515/9781400830244/html

Absil, P.-A., Mahony, R., & Trumpf, J. (2013). An extrinsic look at the Riemannian hessian. In F. Nielsen & F. Barbaresco (Eds.), *Geometric Science of Information* (pp. 361–368). Springer. https://doi.org/10.1007/978-3-642-40020-9_39

Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. Higher-order models. *Journal of Intelligence*, *3*(4), 121–136. https://doi.org/10.3390/jintelligence3040121

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*(5), 676–696. https://doi.org/10.1177/0013164404272507

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry*, *88*(1), 18–27. https://doi.org/10.1016/j.biopsych.2020.01.013

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

Burt, C. (1948). Factor Analysis and canonical correlations. *British Journal of Statistical Psychology*, *1*(2), 95–106. https://doi.org/10.1111/j.2044-8317.1948.tb00229.x

Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika*, *75*(4), 581–612. https://doi.org/10.1007/s11336-010-9178-0

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge Academic. https://www.routledge.com/Statistical-Power-Analysis-for-the-Behavioral-Sciences/Cohen/p/book/9780805802832

Garcia-Garzon, E., Abad, F. J., & Garrido, L. E. (2021). On omega hierarchical estimation: A comparison of exploratory bi-factor analysis algorithms. *Multivariate Behavioral Research*, *56*(1), 101–119. https://doi.org/10.1080/00273171.2020.1736977

Garcia-Garzon, E., Abad, F. J., & Garrido, L. E. (2019). Improving bi-factor exploratory modeling. *Methodology*, *15*(2), 45–55. https://doi.org/10.1027/1614-2241/a000163

Giordano, C., Ones, D. S., Waller, N. G., & Stanek, K. C. (2020). Exploratory bifactor measurement models in vocational behavior research. *Journal of Vocational Behavior*, *120*. https://doi.org/10.1016/j.jvb.2020.103430

Goldberg, L. R. (2006). Doing it all Bass-Ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, *40*(4), 347–358. https://doi.org/10.1016/j.jrp.2006.01.001

Golino, H., & Christensen, A. P. (2022). *EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics.*

Hurley, J. R., & Cattell, R. B. (1962). The procrustes program: Producing direct rotation to test a hypothesized factor structure. *Behavioral Science*, *7*(2), 258–262. https://doi.org/10.1002/bs.3830070216

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory Bi-factor Analysis: The Oblique Case. *Psychometrika*, *77*(3), 442–454. https://doi.org/10.1007/s11336-012-9269-1

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537–549. https://doi.org/10.1007/s11336-011-9218-4

Jimenez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2022). *Dimensionality assessment in bi-factor structures with multiple general factors: A network psychometrics approach.* PsyArXiv. https://doi.org/10.31234/osf.io/2ujdk

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., … Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

Liu, Y. (2020). Riemannian Newton and trust-region algorithms for analytic rotation in exploratory factor analysis. *British Journal of Mathematical and Statistical Psychology.* https://doi.org/10.1111/bmsp.12211

Lorenzo-Seva, U., & Berge, J. M. F. ten. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *2*(2), 57–64. https://doi.org/10.1027/1614-2241.2.2.57

Lorenzo-Seva, U., & Ferrando, P. J. (2019). A general approach for fitting pure exploratory bifactor models. *Multivariate Behavioral Research*, *54*(1), 15–30. https://doi.org/10.1080/00273171.2018.1484339

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*(1), 116–130. https://doi.org/10.3758/s13428-019-01209-1

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*(5), 698–717. https://doi.org/10.1080/00273171.2016.1215898

Maples, J. L., Carter, N. T., Few, L. R., Crego, C., Gore, W. L., Samuel, D. B., Williamson, R. L., Lynam, D. R., Widiger, T. A., Markon, K. E., Krueger, R. F., & Miller, J. D. (2015). Testing whether the DSM-5 personality disorder trait model can be measured with a reduced set of items: An item response theory investigation of the Personality Inventory for DSM-5. *Psychological Assessment*, *27*(4), 1195–1210. https://doi.org/10.1037/pas0000120

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., 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*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, *24*(2), 99–114. https://doi.org/10.1177/01466210022031552

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, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). CA: Muthén & Muthén. https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Myers, N. D., Jin, Y., Ahn, S., Celimli, S., & Zopluoglu, C. (2015). Rotation to a partially specified target matrix in exploratory factor analysis in practice. *Behavior Research Methods*, *47*(2), 494–505. https://doi.org/10.3758/s13428-014-0486-7

Peeters, C. F. W. (2012). Rotational uniqueness conditions under oblique factor correlation metric. *Psychometrika*, *77*(2), 288–292. https://doi.org/10.1007/s11336-012-9259-3

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Moore, T., & Maydeu-Olivares, A. (2011). Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. *Educational and Psychological Measurement*, *71*(4), 684–711. https://doi.org/10.1177/0013164410378690

Revelle, W. (2022). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

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), e0133413. https://doi.org/10.1371/journal.pone.0133413

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

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53–61. https://doi.org/10.1007/BF02289209

Tian, C., & Liu, Y. (2021). A rotation criterion that encourages a hierarchical factor structure. In M. Wiberg, D. Molenaar, J. González, U. Böckenholt, & J.-S. Kim (Eds.), *Quantitative Psychology* (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-030-74772-5_1

Waller, N. G. (2018). Direct Schmid–Leiman transformations and rank-deficient loadings matrices. *Psychometrika*, *83*(4), 858–870. https://doi.org/10.1007/s11336-017-9599-0

Waller, N. G. (2021). *Fungible: Psychometric functions from the waller lab.* [Manual].

Zhang, G., Hattori, M., Trichtinger, L. A., & Wang, X. (2019). Target rotation with both factor loadings and factor correlations. *Psychological Methods*, *24*(3), 390–402. https://doi.org/10.1037/met0000198

Zhang, G., Jiang, G., Hattori, M., & Trichtinger, L. (2020). *EFAutilities: Utility functions for exploratory factor analysis* [Manual]. https://CRAN.R-project.org/package=EFAutilities

Appendix A

The minimization of the extended target criterion (7) can be performed using the gradient projection algorithm (Bernaards & Jennrich, 2005), which is the standard optimization routine implemented in popular statistical software packages such as MPlus (Muthén & Muthén, 2017), lavaan (Rosseel, 2012), psych (Revelle, 2022), and EFAutilities (Zhang et al., 2020). However, the convergence rate of the gradient projection algorithm is linear and may take a long time when the number of factors is large, as usually happens in hierarchical structures. Thus, we decided to implement the Riemannian trust-region method devised by Absil et al. (2007) and outlined in algorithm 2 of Liu (2020). Although the cost of Newton-based routines is more expensive per iteration than that of their gradient-based counterparts, their superlinear rate of convergence (Absil et al., 2007, sec. 4) makes them more suitable for high dimensional settings. Liu (2020) was the first to apply Riemannian Newton algorithms to rotate factor loading matrices and showed a significant speedup in the oblique case when compared to the gradient projection algorithm (Liu, 2020, fig. 4 and 5). To use this optimization routine, an expression of the Riemannian Hessian for the extended target criterion is required.

Define $f_\Lambda(\boldsymbol{\Lambda}) := \|\mathbf{W}_\Lambda \odot (\boldsymbol{\Lambda} - \mathbf{T}_\Lambda)\|^2/2$ and $f_\Phi(\boldsymbol{\Phi}) := \|\mathbf{W}_\Phi \odot (\boldsymbol{\Phi} - \mathbf{T}_\Phi)\|^2/2$ such that the extended target criterion becomes

$$f(\mathbf{X}) = f_\Lambda(\boldsymbol{\Lambda}) + \frac{w}{2} f_\Phi(\boldsymbol{\Phi}), \quad \mathbf{X} \in \mathcal{OB}(q, q). \tag{10}$$

Endowed with the canonical inner product, the set of $q \times q$ normalized columns $\mathcal{OB}(q, q)$ is an embedded Riemannian submanifold of $\mathbb{R}^{q\times q}$ whose tangent space is defined by $\mathcal{T}_\mathbf{X}\mathcal{OB} := \{\mathbf{Z} : \mathrm{ddiag}(\mathbf{X}^\top\mathbf{Z}) = \mathbf{0}\}$ (Absil & Gallivan, 2006, sec. 2). To solve (10) with the gradient projection algorithm of Bernaards and Jennrich (2005), we need to move along the descend direction $-\mathrm{grad} f$, where $\mathrm{grad} f$ is termed the Riemannian gradient of f. Following the notation of Absil and Gallivan (2006), let $\tilde{f}$ be a smooth extension of f to the Euclidean space and let $\mathcal{P}_{\mathcal{T}_\mathbf{X}\mathcal{OB}}\, \mathbf{Z} = \mathbf{Z} - \mathbf{X}\mathrm{ddiag}(\mathbf{X}^\top\mathbf{Z})$ be the projection of $\mathbf{Z}$ onto $\mathcal{T}_\mathbf{X}\mathcal{OB}$ (Absil & Gallivan, 2006, sec. 2). Then, $\text{grad} f(\mathbf{X}) = \mathcal{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{OB}}\, \text{grad}\tilde{f}(\mathbf{X})$ (Absil et al., 2010, equation 3.37), where the Euclidean gradient can be easily found by the chain rule,

$$\text{grad}\tilde{f}(\mathbf{X}) = -\mathbf{X}^{-\top}\text{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})^{\top}\ \mathbf{\Lambda} + w\mathbf{X}\ \text{grad}\tilde{f}_{\Phi}(\mathbf{\Phi}), \tag{11}$$

with $\text{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda}) = \mathbf{W}_{\Lambda}^{2} \odot (\mathbf{\Lambda} - \mathbf{T}_{\Lambda})$ and $\text{grad}\tilde{f}_{\Phi}(\mathbf{\Phi}) = \mathbf{W}_{\Phi}^{2} \odot (\mathbf{\Phi} - \mathbf{T}_{\Phi})$.

For second-order methods, the Riemannian Newton equation becomes $\text{Hess} f(\mathbf{X})[\mathbf{Z}] = -\text{grad} f(\mathbf{X})$ (Absil et al. (2009), equation 6.2), where $\text{Hess} f(\mathbf{X})[\mathbf{Z}]$ is the Riemannian Hessian of f at $\mathbf{X}$ along $\mathbf{Z}$. We may write $\text{Hess} f(\mathbf{X})[\mathbf{Z}]$ as the projection of the directional derivative of the Riemannian gradient along $\mathbf{Z}$ onto $\mathcal{OB}(q,q)$ (Absil et al., 2013, sec. 3):

$$\begin{aligned}\text{Hess} f(\mathbf{X})[\mathbf{Z}] &= \mathcal{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{OB}}\ \mathrm{D}\,\text{grad} f(\mathbf{X})[\mathbf{Z}] \\ &= \mathcal{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{OB}}\ \mathrm{D}\,\text{grad}\tilde{f}(\mathbf{X})[\mathbf{Z}] - \mathbf{Z}\,\text{ddiag}(\mathbf{X}^{\top}\text{grad}\tilde{f}(\mathbf{X})),\end{aligned} \tag{12}$$

This means that we need the directional derivative of the Euclidean gradient along $\mathbf{Z}$. An expression for the directional derivative of the first term of the right-hand part of (11) is given by Liu (2020) as $-\mathbf{X}^{-\top}\mathbf{Z}^{\top}\text{grad}\tilde{f}(\mathbf{X}) + \mathbf{X}^{-\top}\mathrm{D}\,\text{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})[\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top}]^{\top}\mathbf{\Lambda} - \text{grad}\tilde{f}(\mathbf{X})\mathbf{Z}^{\top}\mathbf{X}^{-\top}$ (Appendix A, equation 37). On the other hand, the directional derivative of the second term along $\mathbf{Z}$ is $w\ (\mathbf{Z}\ \text{grad}\tilde{f}_{\Phi}(\mathbf{\Phi}) + \mathbf{X}(\mathbf{W}_{\Phi}^{2} \odot (\mathbf{Z}^{\top}\mathbf{X} + \mathbf{X}^{\top}\mathbf{Z})))$. Thus, the directional derivative for (11) becomes

$$\begin{aligned}\mathrm{D}\,\text{grad}\tilde{f}(\mathbf{X})[\mathbf{Z}] = -\mathbf{X}^{-\top}\mathbf{Z}^{\top}\text{grad}\tilde{f}(\mathbf{X}) + \mathbf{X}^{-\top}\mathrm{D}\,\text{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})[\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top}]^{\top}\mathbf{\Lambda} - \text{grad}\tilde{f}(\mathbf{X})\mathbf{Z}^{\top}\mathbf{X}^{-\top} + \\ w\ (\mathbf{Z}\ \text{grad}\tilde{f}_{\Phi}(\mathbf{X}) + \mathbf{X}(\mathbf{W}_{\Phi}^{2} \odot (\mathbf{Z}^{\top}\mathbf{X} + \mathbf{X}^{\top}\mathbf{Z}))),\end{aligned} \tag{13}$$

where $\mathrm{D}\,\text{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})[\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top}] = \mathbf{W}_{\Lambda}^{2} \odot (\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top})$ for the extended target criterion. An approximate solution for $\mathbf{Z}$ is then found with the truncated conjugate gradient method outlined in algorithm 4 of Liu (2020).

In the same way, to rotate the first and second-order solutions required for the Schmid-Leiman transformation, we need to minimize the quartimin criterion,

$$f(\mathbf{X}) = \frac{1}{4}\|\mathbf{\Lambda}^{2\top}\mathbf{\Lambda}^{2}\mathbf{N}\|^{2}, \quad \mathbf{X} \in \mathcal{OB}(q,q). \tag{14}$$

where $\mathbf{N}$ is a square matrix with zeros on the diagonal and ones elsewhere.

In this case, the Euclidean gradient of a smooth extension of f is

$$\operatorname{grad}\tilde{f}(\mathbf{X}) = -\mathbf{X}^{-\top}\operatorname{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})^{\top}\mathbf{\Lambda}, \tag{15}$$

where $\tilde{f}_{\Lambda}(\mathbf{\Lambda}) = \|\mathbf{\Lambda}^{2\top}\mathbf{\Lambda}^{2}\mathbf{N}\|^{2}/4$ and $\operatorname{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda}) = \mathbf{\Lambda} \odot (\mathbf{\Lambda}^{2}\mathbf{N})$.

Finally, to find the directional derivative of the Euclidean gradient along $\mathbf{Z}$ we simply ignore the last term of the right-hand part of (13), as $\mathbf{\Phi}$ does not affect the quartimin criterion, and replace $\mathrm{D}\operatorname{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})[\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top}]$ with

$$\mathrm{D}\operatorname{grad}\tilde{f}_{\Lambda}(\mathbf{\Lambda})[\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top}] = \mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top} \odot (\mathbf{\Lambda}^{2}\mathbf{N}) + 2\mathbf{\Lambda} \odot ((\mathbf{\Lambda}\mathbf{Z}^{\top}\mathbf{X}^{-\top} \odot \mathbf{\Lambda})\mathbf{N}). \tag{16}$$

Appendix B

Table 4. Estimated loadings for the reduced version of the PID-5-SF with 15 facets or group factors. Loadings with absolute values greater than .20 are shown in bold and underlined. Each facet encompasses 4 items delineated between horizontal bars.

Item	General factors					Group factors														
	G1	G2	G3	G4	G5	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
1	**.73**	-.05	-.04	.12	.02	**.29**	-.05	-.04	.04	-.09	-.02	.00	.01	-.02	.03	.03	.05	.02	.01	.02
2	**.31**	-.01	-.02	**.50**	.07	**.46**	.05	.01	.07	.05	.00	-.04	-.04	-.01	-.06	.03	-.04	-.04	-.06	-.01
3	**.63**	.02	-.04	-.06	.04	**.35**	-.08	-.03	-.08	**-.23**	-.02	.04	-.03	.02	.05	.05	.07	-.04	.03	-.02
4	**.33**	-.03	-.01	**.64**	-.11	**.40**	-.01	.00	-.01	-.04	.01	-.01	.00	-.01	-.03	.00	-.09	.03	-.05	-.03
5	**.59**	.00	.01	-.04	.07	.03	**.54**	.03	.01	-.05	-.01	-.03	.01	-.02	.00	.02	.02	.05	.01	-.04
6	**.76**	.04	-.03	.02	-.10	-.03	**.34**	-.11	.08	.05	.02	.04	.04	.02	.00	-.03	-.03	-.01	.01	-.03
7	**.76**	.07	.00	-.04	-.05	-.02	**.20**	-.10	.14	.05	.02	-.04	-.03	-.02	-.02	-.01	.00	.05	-.03	.06
8	**.60**	.01	.05	.07	.05	.03	**.57**	.06	-.04	.02	-.01	-.01	.01	.00	-.04	-.05	.04	-.04	.00	.00
9	**.59**	-.05	-.01	.00	-.04	-.02	.05	**.63**	.07	.04	-.03	-.01	.00	-.01	.03	.02	.01	.03	.00	.03
10	**.57**	-.02	.04	-.04	-.01	.00	-.02	**.67**	.06	.06	-.01	-.03	-.03	-.03	-.01	.00	.00	.04	-.02	.02
11	**.42**	-.01	.08	.06	.07	.02	.00	**.40**	-.05	-.02	-.06	.00	.08	-.02	-.07	.02	.00	-.10	.00	-.03
12	**.47**	-.01	.02	.03	.01	-.01	.05	**.62**	-.11	-.04	-.04	.04	.04	.05	.00	.02	.02	-.02	-.05	.01
13	.05	**.67**	-.07	.08	-.04	-.01	.04	-.04	**.30**	.02	.01	.04	-.04	.03	-.02	-.04	-.01	-.01	.01	-.12
14	-.04	**.73**	.06	-.05	.03	-.02	.00	-.14	.18	-.02	-.05	-.15	-.04	.02	-.03	-.04	-.02	.13	.00	.04
15	-.09	**.80**	.01	.00	-.05	-.03	.01	-.16	.04	-.11	-.01	.00	.02	.07	.02	.00	-.02	.03	-.01	.02
16	.08	**.71**	.00	.02	.00	.05	.01	-.07	**.22**	-.02	.00	-.03	-.01	.05	.00	-.10	.04	.04	.12	-.04
17	.07	**.43**	.01	.13	-.03	.03	-.06	.02	.05	**.35**	-.03	-.06	.05	-.02	.01	.05	.15	-.04	.07	-.03
18	**.33**	**.55**	-.08	-.12	.00	-.05	.03	-.02	.11	**.39**	.01	-.03	.03	.00	-.02	-.05	.00	.04	.05	-.04
19	**.39**	**.48**	.03	-.03	.03	.04	.02	.02	-.04	**.52**	.00	.04	-.01	.02	.03	-.02	.07	.00	.02	-.03
20	**.31**	**.50**	.04	.07	.01	-.02	-.02	.01	-.09	**.51**	.05	-.04	.01	-.04	.02	-.02	.01	-.02	.00	.05
21	-.02	**.43**	.04	.02	.04	-.02	.03	-.03	.06	.09	**.70**	.02	.01	-.01	-.03	.00	.00	.02	.01	.01
22	.05	**.52**	-.17	-.07	.07	.04	.02	-.06	-.06	-.05	**.30**	-.01	.03	.00	.03	.02	.02	-.04	.02	.06
23	-.02	**.55**	.00	.03	-.03	-.02	-.03	.01	.02	-.01	**.72**	.02	.01	.01	.01	-.01	.01	-.03	.00	-.02
24	-.04	**.56**	.00	.03	.00	.03	-.01	-.07	-.08	-.04	**.59**	-.03	-.05	-.01	-.01	-.01	-.02	.00	.01	.04
25	-.01	-.04	**.59**	.03	.02	-.01	.01	.00	-.01	-.01	-.01	**.55**	-.06	.06	.00	.03	-.01	.11	.04	-.01
26	.02	-.02	**.73**	-.04	-.01	.03	.02	.05	.08	-.03	.04	**.28**	.03	.02	.02	.00	.00	.03	-.03	.13
27	-.05	.00	**.72**	.02	-.04	-.08	.00	-.02	-.07	.02	.01	**.46**	-.07	-.02	.00	.05	-.02	-.03	.02	-.02
28	-.02	.00	**.90**	-.09	-.03	.00	.00	.01	-.06	.00	-.02	.15	.00	-.06	-.07	-.01	-.04	-.03	.03	.04
29	.11	.07	**.52**	.01	.05	.00	.00	.04	-.02	.00	.01	.04	**.40**	.05	.08	.03	-.08	.00	.07	.04
30	-.08	-.06	**.58**	.12	.01	-.07	.00	.05	.01	.05	.01	.05	**.26**	.01	.00	.11	.04	-.01	-.07	.08
31	-.03	-.01	**.88**	.01	-.07	.04	.00	-.01	.05	-.03	.01	**.25**	.05	.00	-.05	.00	-.01	-.03	.01	.11
32	.06	-.05	**.76**	-.01	-.01	-.03	-.01	-.01	-.05	-.02	-.01	-.03	**.44**	-.09	-.03	-.01	.04	-.01	-.03	-.06
33	.02	-.08	**.60**	.00	-.09	-.03	-.05	-.05	.03	.01	.01	-.05	-.04	**.66**	.00	.02	.01	.05	.03	-.03
34	-.05	.00	**.58**	-.06	.00	.02	-.03	-.02	-.05	-.01	-.02	.06	-.06	**.59**	-.01	.00	.03	.09	.10	.00
35	-.03	.04	**.41**	.08	.13	.01	.08	.08	.04	.02	.00	.04	.05	**.49**	-.01	-.04	-.06	-.06	-.04	.08
36	.02	.13	**.50**	-.01	.04	.00	.02	.01	-.02	-.04	.00	.02	.01	**.45**	-.01	-.04	-.06	-.06	-.04	.01

Table 4 (Continuation)

Item	General factors					Group factors														
	G1	G2	G3	G4	G5	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
37	-.03	.04	**.23**	.18	.04	.03	.01	.02	.07	.03	.09	-.07	.04	.04	**.34**	.10	.11	-.07	.00	.03
38	.03	.05	**.49**	.00	.06	.06	.02	.02	.07	.07	-.03	-.05	**.21**	.02	.15	.08	.01	-.16	-.03	.00
39	-.02	-.03	**.28**	**.24**	-.01	-.02	-.04	-.04	-.08	-.03	-.05	.02	-.01	-.03	**.73**	-.01	.02	.01	-.02	.03
40	.10	.03	**.36**	.05	.13	.04	-.03	.03	.17	.03	.02	-.04	.08	-.09	.16	.02	.10	**-.24**	.05	-.07
41	.08	-.01	.03	**.39**	-.04	.01	-.06	.00	-.09	-.03	.01	.01	-.04	-.03	-.02	**.64**	-.03	.08	.00	-.08
42	-.01	-.02	.01	**.48**	-.01	.04	-.05	-.01	-.01	.01	-.01	-.01	.05	.02	.02	**.72**	.02	.00	-.03	.05
43	.01	-.03	.06	**.33**	.05	.03	.04	.01	.06	-.03	-.01	.00	.06	-.01	.03	**.59**	.05	-.04	.00	-.01
44	-.04	-.01	-.02	**.47**	.02	.01	.03	.05	.01	-.01	-.01	.04	-.03	-.01	.03	**.65**	.03	-.05	.03	-.02
45	.03	.04	.02	**.54**	.02	-.02	-.01	.01	.02	.08	.01	.02	.02	.00	.06	.02	**.60**	.00	.03	.01
46	.06	-.12	.00	**.57**	-.06	.04	.02	-.03	.06	-.01	.00	.00	-.01	-.04	.05	.03	**.64**	.00	.04	-.04
47	.01	.07	-.03	**.58**	.00	-.02	.02	.04	-.04	.01	.04	-.02	.00	-.01	.01	.03	**.65**	-.01	-.08	.00
48	.03	.06	-.01	**.53**	.04	.02	.01	.04	-.03	-.04	-.04	-.03	-.01	.00	.00	-.02	**.64**	-.03	.02	-.02
49	.04	-.02	-.07	-.06	**.81**	.02	.01	.02	.05	.02	.00	.15	.05	.02	-.02	.03	-.01	**.42**	.03	-.04
50	-.02	.01	.01	-.04	**.83**	.01	-.01	-.03	.03	.03	-.02	.02	.00	-.01	.03	.01	-.03	**.27**	.05	-.01
51	-.01	-.04	.03	-.07	**.63**	.01	-.09	-.02	-.02	-.07	-.01	**.34**	.10	.19	-.05	-.03	-.02	-.06	-.02	.19
52	.00	-.06	-.04	.17	**.70**	-.12	-.01	-.02	-.02	-.04	-.01	.00	-.05	-.01	-.02	.00	-.10	.01	.11	-.12
53	.00	.02	.05	**.60**	-.05	-.01	-.02	-.04	.03	.03	.00	-.01	-.01	.00	.01	.05	.00	.00	**.62**	.00
54	-.01	.14	.00	**.60**	-.01	-.05	-.04	-.01	.01	-.02	.01	.00	-.03	.05	.01	.03	-.05	.00	**.58**	-.02
55	.04	.00	.06	**.60**	.14	.00	-.05	-.01	-.03	.07	.02	.04	-.05	-.01	-.03	-.08	.04	-.05	**.48**	-.02
56	-.05	.00	.00	**.65**	.09	-.01	.05	-.01	.01	-.01	.01	.05	.10	.01	-.05	-.04	-.03	.06	**.48**	.01
57	-.04	.01	.02	.08	**.46**	**.26**	.11	.06	.00	.08	-.04	.01	.16	.02	.03	.05	.09	.04	**.30**	.06
58	.00	.05	.00	.02	**.66**	-.02	.00	-.03	-.05	.02	.03	-.04	-.02	.02	.02	-.04	-.03	.02	.02	**.56**
59	.02	.03	.06	.02	**.63**	-.01	.00	.06	.04	-.02	.05	.01	-.03	.03	.06	.02	-.04	**-.20**	.00	**.25**
60	.04	.09	.10	.03	**.60**	.19	.06	.01	-.04	-.01	.04	-.03	-.01	.00	.04	-.06	.06	-.01	**.22**	.02

Note. G1 = Negative Affect; G2 = Detachment; G3 = Antagonism; G4 = Disinhibition; G5 = Psychoticism; S1 = Emotional Lability; S2 = Anxiousness; S3 = Separation Insecurity; S4 = Withdrawal, S5 = Anhedonia; S6 = Intimacy Avoidance; S7 = Manipulativeness; S8 = Deceitfulness; S9 = Grandiosity; S10 = Irresponsibility; S11 = Impulsivity; S12 = Distractibility; S13 = Unusual Beliefs; S14 = Eccentricity; S15 = Perceptual Dysregulation.

Table 5. Estimated factor correlations for the PID-5-SF with 15 facets or group factors. Correlations with absolute values greater than .20 are shown in bold and underlined.

Factors	General factors					Group factors														
	G1	G2	G3	G4	G5	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
G1	1.00	**.24**	.12	**.44**	**.34**	.01	-.04	.00	.03	.02	-.05	-.03	.04	-.04	-.02	.01	.10	.01	.00	-.03
G2	**.24**	1.00	**.46**	**.47**	**.44**	.00	.05	-.11	.00	-.01	-.01	-.04	-.01	.06	.01	-.08	.02	-.01	.07	.05
G3	12	**.46**	1.00	**.50**	**.44**	-.02	-.01	.09	-.04	.01	.01	.00	-.02	-.02	-.03	.07	-.01	.00	.05	.04
G4	**.44**	**.47**	**.50**	1.00	**.64**	.00	.00	.01	.02	.03	.03	.02	.01	-.03	.01	.04	-.03	-.04	-.06	-.05
G5	**.34**	**.44**	**.44**	**.64**	1.00	.04	.04	.05	-.01	-.01	.02	-.01	.05	.03	.04	-.02	.01	-.04	.07	.01
S1	.01	.00	-.02	.00	.04	1.00	.02	.00	.01	-.01	-.01	-.02	-.01	.00	.00	.05	.03	-.02	-.02	-.01
S2	-.04	.05	-.01	.00	.04	.02	1.00	.07	.03	.02	.00	-.01	.02	.00	-.03	-.03	.03	.01	-.01	.00
S3	.00	-.11	.09	.01	.05	.00	.07	1.00	.00	.04	-.07	.01	.04	.01	-.01	.03	.04	-.03	-.05	.02
S4	.03	.00	-.04	.02	-.01	.01	.03	.00	1.00	.04	.01	-.02	.01	.01	-.01	.00	.01	.02	.01	-.01
S5	.02	-.01	.01	.03	-.01	-.01	.02	.04	.04	1.00	.03	-.03	.02	-.02	.01	-.03	.03	.00	.05	.00
S6	-.05	-.01	.01	.03	.02	-.01	.00	-.07	.01	.03	1.00	.00	-.01	.00	.00	-.01	.01	-.02	.01	.03
S7	-.03	-.04	.00	.02	-.01	-.02	-.01	.01	-.02	-.03	.00	1.00	.01	.04	-.01	.01	-.03	.02	.03	-.01
S8	.04	-.01	-.02	.01	.05	-.01	.02	.04	.01	.02	-.01	.01	1.00	-.02	.01	.03	.01	-.03	-.01	.02
S9	-.04	.06	-.02	-.03	.03	.00	.00	.01	.01	-.02	.00	.04	-.02	1.00	-.01	-.01	-.04	.02	.03	.03
S10	-.02	.01	-.03	.01	.04	.00	.00	.01	.01	-.02	.00	-.01	.01	-.01	1.00	.06	.08	.00	.01	.03
S11	.01	-.08	.07	.04	-.02	.05	-.03	.03	.00	-.03	-.01	.01	.03	-.01	.06	1.00	.05	.00	.00	-.02
S12	.10	.02	-.01	-.03	.01	.03	.0	.04	.01	.03	.01	-.03	.01	-.04	.08	.05	1.00	-.04	.03	-.03
S13	.01	-.01	.00	-.04	-.04	-.02	.01	-.03	.02	.00	-.02	.02	-.03	.02	.00	.00	-.04	1.00	.03	-.01
S14	.00	.07	.05	-.06	.07	-.02	-.01	-.05	.01	.05	.01	.03	-.01	.03	.01	.00	.03	.03	1.00	-.02
S15	-.03	.05	.04	-.05	.01	-.01	.00	.02	-.01	.00	.03	-.01	.02	.03	.03	-.02	-.03	-.01	-.02	1.00

Note. G1 = Negative Affect; G2 = Detachment; G3 = Antagonism; G4 = Disinhibition; G5 = Psychoticism; S1 = Emotional Lability; S2 = Anxiousness; S3 = Separation Insecurity; S4 = Withdrawal, S5 = Anhedonia; S6 = Intimacy Avoidance; S7 = Manipulativeness; S8 = Deceitfulness; S9 = Grandiosity; S10 = Irresponsibility, S11 = Impulsivity; S12 = Distractibility; S13 = Unusual Beliefs; S14 = Eccentricity; S15 = Perceptual Dysregulation.

Tables

Table 1. A random sample of simulated parameters under each of the IC, ICB, ICP and ICBP structures. In the ICB and ICBP structures every pair of group factors belonging to the same general factor shares one indicator cross-loading while in the ICP and ICBP structures one item per group factor only loads on the general factor.

Item	IC									ICB								
	G1	G2	S1	S2	S3	S4	S5	S6	$\mathbf{h^2}$	G1	G2	S1	S2	S3	S4	S5	S6	$\mathbf{h^2}$
1	.45		.60						.57	.35		.53		.40				.57
2	.47		.50						.48	.47		.50						.48
3	.51		.40						.42	.51		.40						.42
4	.58			.60					.70	.51		.40	.53					.70
5	.44			.50					.44	.44			.50					.44
6	.58			.40					.50	.58			.40					.50
7	.59				.60				.71	.52			.40	.53				.71
8	.53				.50				.53	.53				.50				.53
9	.53				.40				.44	.53				.40				.44
10		.41				.60			.53		.30				.53		.40	.53
11		.44				.50			.44		.44				.50			.44
12		.44				.40			.35		.44				.40			.35
13		.54					.60		.65		.46				.40	.53		.65
14		.48					.50		.48		.48					.50		.48
15		.55					.40		.47		.55					.40		.47
16		.50						.60	.61		.41					.40	.53	.61
17		.54						.50	.55		.54						.50	.55
18		.60						.40	.52		.60						.40	.52
Avg									.52									.52

Item	ICP									ICBP								
	G1	G2	S1	S2	S3	S4	S5	S6	$\mathbf{h^2}$	G1	G2	S1	S2	S3	S4	S5	S6	$\mathbf{h^2}$
1	.45		.60						.57	.35		.53		.40				.57
2	.69		.01						.48	.69		.01						.48
3	.51		.40						.42	.51		.40						.42
4	.58			.60					.70	.51		.40	.53					.70
5	.67			.01					.44	.67			.01					.44
6	.58			.40					.50	.58			.40					.50
7	.59				.60				.71	.52			.40	.53				.71
8	.73				.01				.53	.73				.01				.53
9	.53				.40				.44	.53				.40				.44
10		.41				.60			.53		.30				.53		.40	.53
11		.67				.01			.44		.67				.01			.44
12		.44				.40			.35		.44				.40			.35
13		.54					.60		.65		.46				.40	.53		.65
14		.69					.01		.48		.69					.01		.48
15		.55					.40		.47		.55					.40		.47
16		.50						.60	.61		.41					.40	.53	.61
17		.74						.01	.55		.74						.01	.55
18		.60						.40	.52		.60						.40	.52
Avg									.52									.52

Note. The **Avg** row is for the average communality, $\mathbf{h^2}$.

Table 2. Marginal outcomes for each variable level, structure and method. Marginal average congruence coefficients (ACC) equal or greater than .95 and marginal root-mean square residuals of the general factor correlations ($\hat{\mathbf{\Phi}}_g$ RMSE) equal or smaller than .05 are shown in bold and underlined.

Variable	ACC		$\hat{\mathbf{\Phi}}_g$ RMSE	
	SL	GSLiD	SL	GSLiD
N.GF				
2	.943	**.970**	.060	**.029**
3	.935	**.961**	.076	**.038**
4	.927	**.953**	.090	**.044**
5	.920	.945	.102	**.049**
COR.GF				
no	.931	**.956**	**.031**	**.022**
yes	.931	**.958**	.133	.058
N				
500	.911	.934	.100	.053
1000	.935	**.962**	.080	**.039**
2000	.947	**.976**	.065	**.028**
VAR.GRF				
4	.909	.944	.084	**.039**
5	.933	**.961**	.082	**.040**
6	**.951**	**.967**	.079	**.041**
NUM.GRF				
4	.935	**.959**	.080	**.042**
5	.932	**.959**	.082	**.040**
6	.926	**.954**	.084	**.038**
CROSS.GRF				
no	**.955**	**.962**	.068	**.040**
yes	.907	**.953**	.095	**.039**
LOAD.GRF				
low	.902	.923	.087	**.040**
medium	.937	**.966**	.081	**.040**
high	**.954**	**.983**	.077	**.039**
LOAD.GF				
low	.919	.947	.088	**.050**
medium	.932	**.957**	.081	**.039**
high	.942	**.967**	.077	**.031**
PURE				
no	**.966**	**.966**	.084	**.049**
yes	.896	.948	.080	**.031**
STRUCTURES				
IC	**.975**	**.965**	.069	**.050**
ICB	**.957**	**.968**	.098	**.048**
ICP	.935	**.959**	.068	**.031**
ICBP	.857	.938	.092	**.031**
TOTAL	.931	**.957**	.082	**.040**

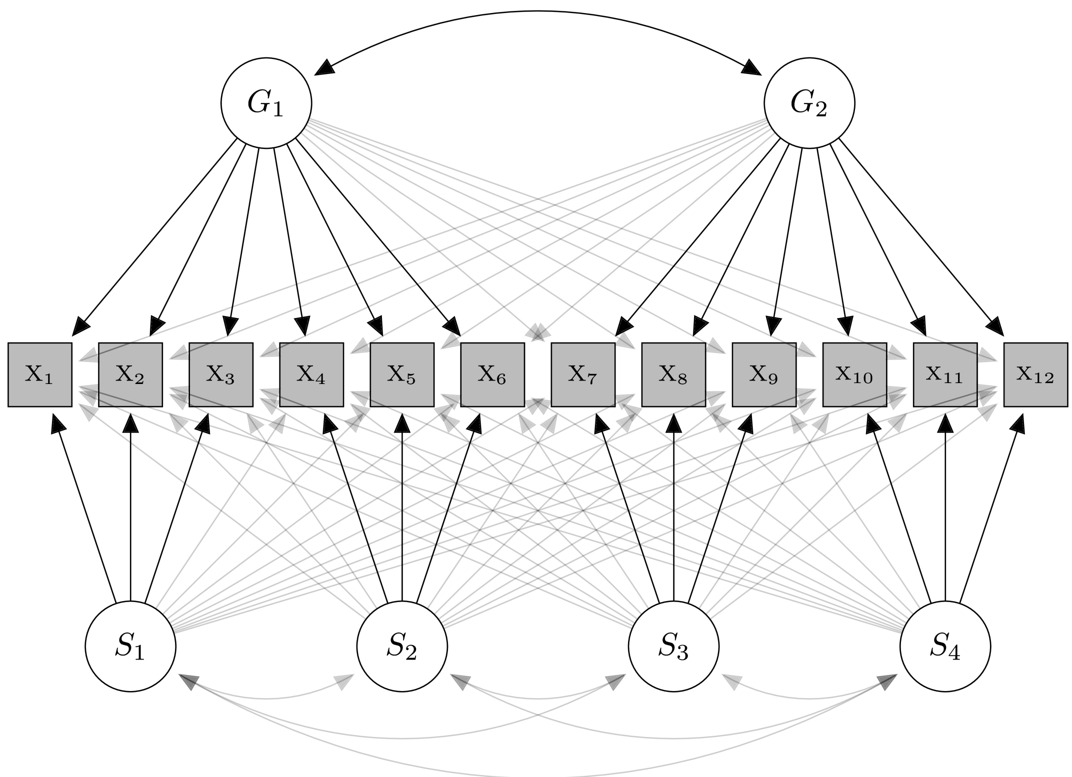
Note. N.GF = number of general factors; COR.GF = Correlation between general factors; N = sample size; VAR.GRF = Number of indicators per group factor; NUM.GRF = number of group factors per general factor; CROSS.GRF = cross-loadings in the group factors; LOAD.GF = loadings on the general factors; LOAD.GRF = loadings on the group factors; PURE.GF = pure indicators of the general factors; IC = Independent cluster structure: neither cross-loadings nor pure indicators are present; ICB = Independent cluster basis: cross-loadings but not pure indicators are present; ICP = Independent cluster pure: pure indicators but not cross-loadings are present; ICBP = Independent cluster pure basis: both cross-loadings and pure indicators are present.

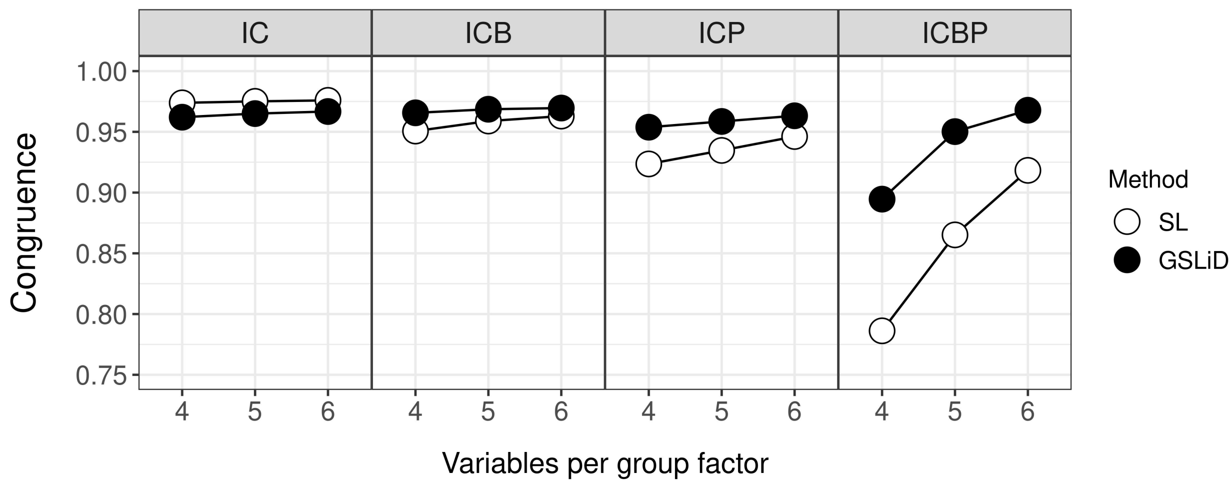
Table 3. Partial omega squared coefficients (Ω^2_{prtl}) from the ANOVAs on the average congruence coefficients (ACC) and the root-mean square residuals of the general factor correlations ($\hat{\mathbf{\Phi}}_g$ RMSE) for all the 9 main effects, and for the remaining coefficients whose $\Omega^2_{prtl} > .14$ in at least one method.

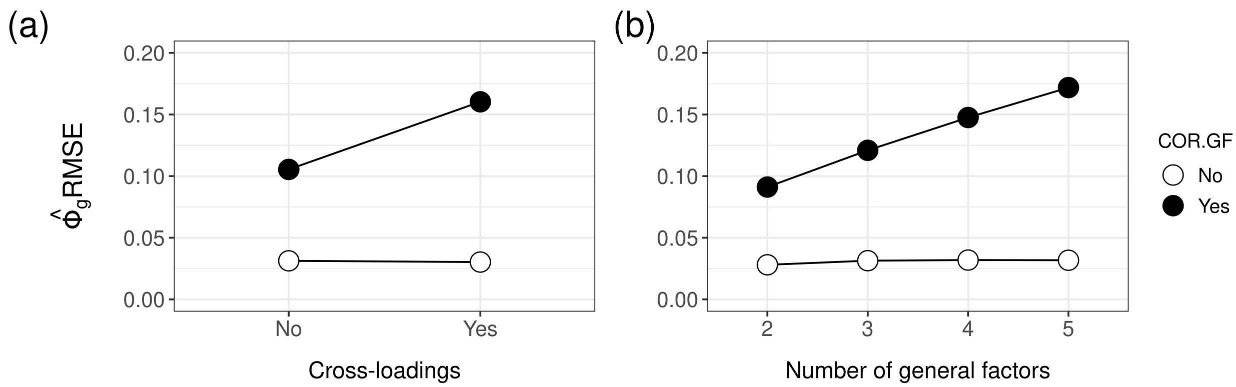
Coefficients	ACC		$\hat{\mathbf{\Phi}}_g$ RMSE	
	SL	GSLiD	SL	GSLiD
Main effects				
N.GF	.34	.27	.39	.17
COR.GF	.00	.00	.87	.55
N	.60	.57	.35	.29
VAR.GRF	.68	.28	.01	.00
NUM.GRF	.10	.02	.01	.01
CROSS.GRF	.80	.08	.32	.00
LOAD.GRF	.77	.73	.05	.00
LOAD.GF	.37	.23	.05	.19
PURE.GF	.90	.26	.01	.23
Two-way interactions				
CROSS.GRF × PURE.GF	.62	.14	.00	.00
CROSS.GRF × VAR.GRF	.52	.17	.00	.00
CROSS.GRF × LOAD.GRF	.17	.04	.03	.00
PURE.GF × LOAD.GRF	.17	.18	.01	.02
PURE.GF × VAR.GRF	.59	.21	.00	.00
N × LOAD.GRF	.31	.32	.01	.00
COR.GF × N.GF	.00	.00	.35	.12
COR.GF × CROSS.GRF	.00	.01	.34	.00
COR.GF × PURE.GF	.01	.01	.00	.28
COR.GF × LOAD.GF	.00	.00	.01	.24
Three-way interactions				
CROSS.GRF × PURE.GF × VAR.GRF	.42	.17	.00	.00

Note. N.GF = number of general factors; COR.GF = Correlation between general factors; N = sample size; VAR.GRF = Number of indicators per group factor; NUM.GRF = number of group factors per general factor; CROSS.GRF = cross-loadings in the group factors; LOAD.GF = loadings on the general factors; LOAD.GRF = loadings on the group factors; PURE.GF = pure indicators of the general factors.

Figures


G_1
G_2
X_1
X_2
X_3
X_4
X_5
X_6
X_7
X_8
X_9
X_{10}
X_{11}
X_{12}
S_1
S_2
S_3
S_4


IC
ICB
ICP
ICBP
Congruence
1.00
0.95
0.90
0.85
0.80
0.75
4
5
6
Variables per group factor
Method
SL
GSLiD


(a)
(b)
$\hat{\Phi}_g$RMSE
0.20
0.15
0.10
0.05
0.00
No
Yes
Cross-loadings
2
3
4
5
Number of general factors
COR.GF
No
Yes

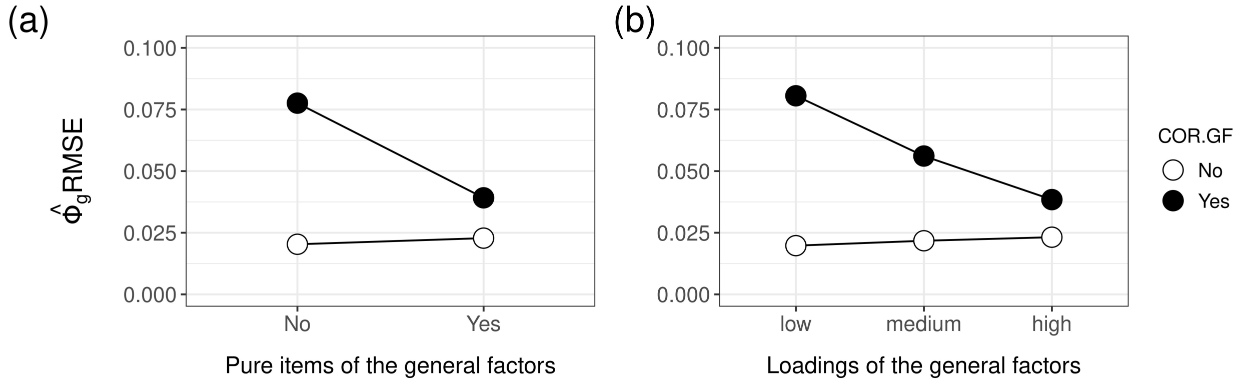

(a)
(b)
$\hat{\Phi}_g$RMSE
0.100
0.075
0.050
0.025
0.000
No
Yes
Pure items of the general factors
low
medium
high
Loadings of the general factors
COR.GF
No
Yes

Figure captions

Figure 1. Illustration of an exploratory bi-factor model with two general factors (G) and four group factors (S) for twelve indicators (X). Dark arrows correspond to salient loadings and light arrows indicate possible cross-loadings and correlations.

Figure 2. Interaction PURE.GF $\times$ CROSS.GRF $\times$ VAR.GRF on the ACC for GSLiD and SL.

Figure 3. Interactions COR.GF $\times$ CROSS.GRF (a) and COR.GF $\times$ N.GF (b) on the $\hat{\boldsymbol{\Phi}}_g$ RMSE for SL.

Figure 4. Interactions COR.GF $\times$ PURE.GF (a) and COR.GF $\times$ LOAD.GF (b) on the $\hat{\boldsymbol{\Phi}}_g$ RMSE for GSLiD.