

Recovering Bifactor Models: A Comparison of Seven Methods

Casey Giordano and Niels G. Waller
University of Minnesota—Twin Cities

Abstract

The last decade has witnessed a resurgence of interest in exploratory bifactor analysis models and the concomitant development of new methods to estimate these models. Understandably, due to the rapid pace of developments in this area, existing Monte Carlo comparisons of bifactor analysis have not included the newest methods. To address this issue, we compared the model recovery capabilities of 5 existing methods and 2 newer methods (Waller, 2018a) for exploratory bifactor analysis. Our study expands upon previous work in this area by comparing (a) a greater number of estimation algorithms and (b) by including both nonhierarchical and hierarchical bifactor models in our study design. In aggregate, we conducted almost 3 million exploratory bifactor analyses to identify the most accurate methods. Our results showed that, when compared with the alternatives, the rank-deficient Schmid-Leiman and Direct Schmid-Leiman methods were better able to recover both nonhierarchical and hierarchical bifactor structures.

Translational Abstract

The last decade has witnessed a resurgence of interest in exploratory bifactor analysis models and the concomitant development of new methods to estimate these models. Understandably, due to the rapid pace of developments in this area, existing comparisons of bifactor analysis routines have not included the newest methods. To address this issue, we compared the model recovery capabilities of five existing and two newer methods (Waller, 2018a) for exploratory bifactor analysis. Our study expands upon previous work in this area by comparing (a) a greater number of estimation algorithms and (b) by including both nonhierarchical and hierarchical bifactor models in our study design. In aggregate, we conducted almost three million exploratory bifactor analyses to identify the most accurate methods.

Keywords: bifactor analysis, Direct Schmid-Leiman, Schmid-Leiman, Monte Carlo simulation

Supplemental materials: <http://dx.doi.org/10.1037/met0000227.supp>

After years of relative neglect, bifactor analysis (Holzinger & Swineford, 1937; Schmid & Leiman, 1957) has resurged as a popular method for exploring the latent structure of multidimensional data (Mansolf & Reise, 2016; Reise, 2012).¹ Renewed interest in the bifactor model is due, in part, to recent developments of exploratory (Jennrich & Bentler, 2011, 2012) and target bifactor rotation methods (Abad, Garcia-Garzon, Garrido, & Barrada, 2017; Reise, Moore, & Maydeu-Olivares, 2011). These methods build upon foundational work by Holzinger and Swineford (1937) and Schmid and Leiman (1957; see also Wherry, 1959) by elucidating the hierarchical and nonhierarchical latent structures of multidimensional item pools. In this article, we com-

pare the relative performance of seven methods for exploratory bifactor analysis in recovering known bifactor structures with various psychometric properties. An overarching goal of our work is to provide empirically supported guidance to researchers who wish to estimate bifactor models in their research.

Our article is structured as follows. First, we briefly review the history of bifactor measurement models and focus our comments on exploratory bifactor analysis (EBF). Next, we summarize the results of a comprehensive Monte Carlo study in which we compared the following EBF methods: (a) the traditional Schmid-Leiman procedure (SL; Schmid & Leiman, 1957); (b) a recently proposed extension of the SL called the Direct Schmid-Leiman method (DSL; Waller, 2018a); (c) a recently proposed Direct Bifactor method (DBF; Waller, 2018a); (d) a partially specified target bifactor method first described by Reise, Moore, and Haviland (2010; see also Moore, Reise, Depaoli, & Haviland, 2015) and later (e) modified by Abad et al. (2017); (f) an orthogonal full-rank bifactor rotation method by Jennrich and Bentler (2011, 2013); and (g) an oblique full-rank bifactor rotation method by the same authors (Jennrich & Bentler, 2012). In this design we varied

This article was published Online First July 25, 2019.

Casey Giordano and Niels G. Waller, Department of Psychology, University of Minnesota—Twin Cities.

Portions of this research were presented at the (2018) 33rd annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL and the (2017) annual meeting of the Society for Multivariate Behavioral Psychology, Minneapolis, MN.

Correspondence concerning this article should be addressed to Niels G. Waller, Department of Psychology, University of Minnesota—Twin Cities, 75 East River Road, Minneapolis, MN 55455. E-mail: nwaller@umn.edu

¹ The online supplemental material includes graphical summaries of the annual and cumulative citation counts for the classic Schmid and Leiman (1957) article.

(a) sample size, (b) number of variables per group factor, (c) number of group factors, (d) the presence of factor cross-loadings, (e) the magnitude of the general factor loadings, (f) the magnitude of the group factor loadings, and (g) the model structure (hierarchical vs. nonhierarchical factor models). This Monte Carlo design was used in two simulation studies. In the first study, we compared all of the EBF methods in 48,600 data sets representing 50 replicates for each of the 972 combinations of the aforementioned fully crossed factorial design. In the second study, we supplemented the best performing methods from Study 1 with two additional variations of the DSL and DBF methods.² These analyses were performed on 486,000 new data sets representing 500 replicates for the 972 design combinations. In aggregate, we conducted 2,818,800 exploratory bifactor analyses in the two Monte Carlo simulations.

A Brief History of Exploratory Bifactor Analysis

As noted by several authors (Beaujean, 2015; Harman, 1954; Holzinger, 1936, 1945), the bifactor model was originally developed by Karl Holzinger while collaborating with Charles Spearman on the Unitary Traits Study (Holzinger, 1936, 1945; Spearman & Holzinger, 1934). Holzinger's bifactor model (Holzinger & Swineford, 1937) grew out of Spearman's (1904) development of the factor analysis model. In his initial report on factor analysis, Spearman (1904) noticed that the academic grades of British schoolchildren were positively correlated. To account for these correlations, Spearman proposed the so-called two-factor model³ of mental abilities in which all variables share variance with a general intelligence dimension that is commonly labeled *g* (see Spearman, 1927). Soon after Spearman (1927) published his magnum opus, *The Abilities of Man*, two lines of criticism concerning his model of general intelligence appeared in the literature. In the first line of criticism, several reviewers (e.g., Camp, 1932; Piaggio, 1931; Wilson, 1928) noted that an important aspect of Spearman's model—to wit, the estimation of factor scores—was mathematically unidentified, thus making it impossible to uniquely calculate a person's standing on *g* (for a readable discussion of this issue, see Steiger, 1979). In the second line of criticism, L. L. Thurstone and others (e.g., Garnett, 1919) suggested that Spearman's model of general intelligence “is inadequate for the multi-dimensionality of the mental abilities” (Thurstone, 1934, p. 2). As an alternative to Spearman's model, Thurstone (1947) proposed a hierarchical model that gained favor in the psychometrics community (Reise, 2012; Thurstone, 1934, 1935, 1947; Tucker, 1940) because it included multiple (correlated) common factors. Nevertheless, not all researchers were willing to abandon models with a broad general factor. To accommodate these individuals, Holzinger and Spearman proposed the so-called bifactor measurement model (Holzinger, 1936; Holzinger & Swineford, 1937; Spearman & Holzinger, 1934) that includes (a) a general factor on which all items load and (b) one or more group factors on which subsets of items load. In this hybrid model, the general and group factors are mutually uncorrelated. Unfortunately, due to the overwhelming popularity of Thurstone's work in this area, Holzinger's bifactor model was not widely embraced by the psychometrics community.

Although Holzinger's model failed to settle the debate between members of the Spearman and Thurstone camps, his model influenced later developments on the hierarchical representation of latent

factors. Among these developments, Cox (1939) demonstrated (with simulated data) that a test battery influenced by a general factor and three uncorrelated group factors could be well-represented by Thurstone's correlated factors model. Thurstone and Thurstone (1941) later showed that a correlated factors model can itself be factor analyzed into a second-order model. Combining these ideas, Thomson (1946/1951) described a prototypical bifactor model that allowed researchers to convert a correlated factors solution into an orthogonal model with a general dimension and several group factors (see also, Thurstone, 1947, Chapter 18). Expanding on this work, Schmid and Leiman (1957, see also Wherry, 1959) later developed their well-known variant of the exploratory bifactor model in which any higher-order factor solution can be transformed into a solution with mutually orthogonal factors. Commenting on these developments, Carroll (1993) noted that the Schmid and Leiman and Thomson procedures were never widely used, in part, because computer routines for these methods were never incorporated into popular statistical software. For this and other reasons, Holzinger's bifactor model laid dormant for several decades.

More recently, there has been a rediscovery of bifactor measurement models (see Reise, 2012). Dovetailing this renewed interest (e.g., Canivez, Watkins, & Dombrowski, 2016; Dombrowski, 2014; Revelle & Wilt, 2013; Sharp et al., 2015), psychometricians have developed new methods for conducting exploratory⁴ bifactor analysis (e.g., Jennrich & Bentler, 2011, 2012, 2013; Reise, Moore, & Haviland, 2010; Schmid & Leiman, 1957; Waller, 2018a). In the next section, we describe the mathematics and underlying logic of the seven EBF methods that were included in our Monte Carlo study. We begin our review by describing the Schmid-Leiman method (Schmid & Leiman, 1957).

Seven Methods for Exploratory Bifactor Analysis

The Schmid-Leiman Method

The Schmid-Leiman (SL) method is “a procedure for transforming an oblique factor analysis solution containing a hierarchy of higher-order factors into an orthogonal solution” (Schmid & Leiman, 1957, p. 53) that includes one or more general factors and two or more group factors.⁵ The underlying mathematics of the two-level SL solution can be described as follows. Let

² Specifically, for the EBF methods that use target rotations, we used the population bifactor loadings matrices—and dichotomized versions of these matrices consisting of signed ones and zeros—as the targets.

³ In Spearman's (1927) two-factor model of general intelligence, each observed variable putatively shares variance with two reliable latent factors: (a) a common factor that influences all observed variables and (b) a specific factor that influences only one observed variable.

⁴ For discussions about the properties of bifactor models in relation to hierarchical models, see Mulaik and Quartetti (1997) and Yung, Thissen, & McLeod, 1999. Treatments on confirmatory bifactor analysis are discussed by Murray and Johnson (2013), Gignac (2008, 2016), and Mansolf and Reise (2017).

⁵ Technically, the Schmid-Leiman procedures requires three or more group factors at each level of the hierarchy to be fully identified, although it is possible to obtain parameter estimates when only two group factors are present. Although in their original description of the method, the authors demonstrated this method on a three-level model, most contemporary applications of the SL approach have focused on two-level models only.

$$\Sigma = \Lambda_1 \Phi_1 \Lambda_1^T + \Theta_1^2, \quad (1)$$

where Σ represents a $p \times p$ population correlation matrix for the observed variables, Λ_1 denotes a $p \times k$ matrix of obliquely rotated common factor loadings, Φ_1 denotes a $k \times k$ matrix of common factor intercorrelations, and Θ_1^2 is a $p \times p$ diagonal matrix of the uniqueness factor variances. For all terms, we use the subscript 1 to denote parameter matrices in the first-order solution of a multi-order factor analysis model. As noted by Schmid and Leiman (1957), when Φ_1 is not an identity matrix (and $k \geq 3$) then it is possible to factor analyze the lower-order factor correlations (in Φ_1) to extract broader factors at a higher level of analysis. For example, if Φ_1 can be represented by a single factor model then

$$\Phi_1 = \lambda_2 \lambda_2^T + \Theta_2^2, \quad (2)$$

where λ_2 is a $k \times 1$ vector of loadings of the lower-order factors on a higher-order factor, and Θ_2^2 is a $k \times k$ diagonal matrix of uniqueness factor variances. For all terms in (2), we use the subscript 2 to denote the second-order solution. Notice that by substituting (2) into (1), we produce a second-order factor model,

$$\Sigma = \Lambda_1 (\lambda_2 \lambda_2^T + \Theta_2^2) \Lambda_1^T + \Theta_1^2, \quad (3)$$

where all terms are as previously defined. Schmid and Leiman (1957) described a method for transforming the full model in (3) into a hierarchical bifactor analysis solution. In this method, the authors noted that one can obtain loadings of the observed variables on a general factor (\mathbf{g}) by multiplying the loadings of the first- and second-order factors from (3), such that

$$\mathbf{g} = \Lambda_1 \lambda_2. \quad (4)$$

In modern parlance, the general factor loadings (\mathbf{g}) represent the direct effects of the general factor on the observed variables (Yung, Thissen, & McLeod, 1999). The Schmid and Leiman (1957) method also provides a means for estimating the observed variable loadings on the group factors. To estimate these loadings, one multiplies the first-order factor loadings (Λ_1) from (1) by the second-order uniqueness standard deviations (Θ_2) from (3),

$$\mathbf{D} = \Lambda_1 \Theta_2, \quad (5)$$

where \mathbf{D} is a $p \times k$ matrix of indicator (observed variable) loadings on the group factors. Assembling the results of (4) and (5) into a single matrix, \mathbf{B}_{SL} , the full $p \times (k + 1)$ SL solution of general and group factor loadings equals,

$$\mathbf{B}_{SL} = [\mathbf{g} : \mathbf{D}]. \quad (6)$$

As noted by several authors (e.g., Waller, 2018a; Yung et al., 1999), \mathbf{B}_{SL} is a rank-deficient matrix in which the general factor loadings are linear combinations of the (k) group factor loadings. Due to these constraints, \mathbf{B}_{SL} —a matrix of order $p \times (k + 1)$ —has rank k . The rank-deficiency of \mathbf{B}_{SL} imposes proportionality constraints on the SL factor loadings (Waller, 2018a; Yung et al., 1999).

Two Direct Schmid-Leiman Methods

Extending the Schmid-Leiman method, Waller (2018a) recently described a general procedure for conducting exploratory bifactor analysis that he called the Direct Schmid-Leiman approach. In this work, Waller (2018a) noted that for any factor model with three or

more common factors, there are an infinite number of SL transformations that fit the data equally well (Waller, 2018a) because of the rotational indeterminacy of the factor analysis model. From this observation, Waller wondered whether it was possible to find the best SL solution from among the set of infinite solutions. Here, “best” is defined as the solution with the smallest root mean squared error (RMSE) between the population and estimated bifactor factor loadings matrices. Waller (2018a) proved that under ideal conditions, a best-fitting solution could be found by performing an orthogonal Procrustes rotation (Cliff, 1966; Kristof, 1970; Schönemann, 1966) of an augmented factor loadings matrix, $\tilde{\mathbf{A}}$, such that,

$$\operatorname{argmin} f(\mathbf{T}) = \|\tilde{\mathbf{A}} \mathbf{T}^T - \mathbf{B}_{pop}\|_F^2, \quad (7)$$

where $\tilde{\mathbf{A}}$ is a $p \times (k + 1)$ augmented factor structure matrix that is constructed by appending a column of zeros to a $p \times k$ matrix of unrotated factor loadings,⁶ \mathbf{T} is a $(k + 1) \times (k + 1)$ orthogonal rotation matrix, \mathbf{B}_{pop} is the $p \times (k + 1)$ matrix of population bifactor loadings, $\|\cdot\|_F$ denotes the Frobenius (i.e., the generalized Euclidean) norm operator,⁷ and argmin denotes the argument of the minimum (in this case we seek a \mathbf{T} that minimizes the function). Note that \mathbf{T}^T is the orthogonal rotation matrix that minimizes the squared distance between $\tilde{\mathbf{A}}$ and \mathbf{B}_{pop} . Methods for finding \mathbf{T} have been described by several authors (Cliff, 1966; Schönemann, 1966; Waller, 2018a, 2018c).

From a practical standpoint, minimizing (7) is hampered by the fact that population loadings are not known in empirical applications (although they are known in simulation work). To address this issue, Waller (2018a) described two procedures for estimating \mathbf{B}_{pop} , both of which require the investigator to generate a target matrix that approximates \mathbf{B}_{pop} . The first method requires an investigator to specify a target matrix of salient and nonsalient factor loadings (both on the group and general factors). In this hypothesis-driven approach, elements of the target matrix are assigned a value of ± 1 when an indicator is a putative marker of the factor under consideration. Nonsalient loadings are assigned a value of 0. This user-specified target replaces \mathbf{B}_{pop} in the orthogonal Procrustes rotation.

Waller’s second method for generating a target matrix is more data driven and is thus more exploratory. In this method, the data determine the pattern of (signed) ones and zeros. This is accomplished as follows. First, a researcher extracts and obliquely rotates a factor analysis model with k factors. Next, using the factor pattern from this model, the matrix of common factor loadings is converted into a matrix of (signed) ones and zeros by dichotomizing the analogue loadings at a user-defined threshold (e.g., $\lambda_{ij} = |.25|$). A unity vector (i.e., a vector of all ones signifying that all indicators load on the general factor) is appended to this signed binary matrix to create a $p \times (k + 1)$ empirically generated target matrix. Notice that in both methods, this seemingly crude approach for approximating \mathbf{B}_{pop} (with binary values) can be justified by recalling that in moderate to small samples, unit weights often

⁶ Waller (2018a) appends a column of zeros to the unrotated factor loadings matrix, \mathbf{A} , to create a matrix $\tilde{\mathbf{A}}$ that has a column order equal to that of \mathbf{B}_{pop} (i.e., $k + 1$) and a matrix rank that is equal to the rank of \mathbf{A} (i.e., k).

⁷ In the current context, the Frobenius norm operator can be defined as follows. Let $\mathbf{C} = \tilde{\mathbf{A}} \mathbf{T}^T - \mathbf{B}_{pop}$. Then, $\|\mathbf{C}\|_F = \sqrt{\sum_{i=1}^p \sum_{j=1}^{k+1} c_{ij}^2}$.

outperform other weighting schemes when considering predictive accuracy (cf. Wainer, 1976, 1978; Wilks, 1938). We call solutions obtained by these methods a Direct Schmid-Leiman (DSL) solution. It merits comment that, like the original Schmid-Leiman approach, the DSL was designed to estimate the parameters of a hierarchical bifactor analysis model.

In this same article, Waller (2018a) also described a method for estimating the parameters of nonhierarchical bifactor models using orthogonal Procrustes rotations. In this method, rather than rotating an augmented and rank-deficient factor solution (of column order $k + 1$) to a target matrix, he noted that one could rotate a full-rank loadings matrix (of column order $k + 1$) to a target matrix via an orthogonal Procrustes transformation. We call this approach a Direct Bifactor (DBF) method. Note that the aforementioned methods for generating a DSL target matrix can also be used to generate a DBF target matrix.

Estimating Bifactor Models Using Partially Specified Target Rotations

We now describe two methods for estimating nonhierarchical bifactor models that rely on Browne's algorithm (Browne, 1972, 2001; see also Horst, 1941) for rotating a matrix to a partially specified target. Reise et al. (2010) were the first to apply Browne's algorithm to estimate a bifactor model. Abad et al. (2017) called this method a Schmid-Leiman target (SLt) rotation. We use the same nomenclature. In a later development, Moore et al. (2015) suggested that Browne's method could be improved by applying it in an iterative manner. Abad et al. (2017) used the iterative SLt to estimate nonhierarchical bifactor models and called their method SLi (to acknowledge its ties to the Schmid-Leiman approach). In what follows, we describe the underlying mathematics of the SLt and SLi for estimating bifactor loadings matrices.

For the SLt, Reise et al. (2010) suggested that a partially specified (PS) target matrix for \mathbf{B}_{pop} (\mathbf{B}_{PS} ; Browne, 1972, 2001; Horst, 1941) could be created as follows. Using the solution from a previously estimated SL model (a) fix the nonsalient elements of the loadings matrix (i.e., loadings [in absolute value] below a user-defined threshold) to zero and (b) allow all other loadings to be freely estimated. Next, use Browne's (1972, 2001) algorithm to rotate a $k + 1$ factor loadings matrix to minimize,

$$f(\mathbf{B}_{\text{PS}}) = \sum_{i=1}^p \sum_{j=1}^{k+1} w_{ij} (\lambda_{ij} - b_{\text{PS}(ij)})^2, \quad (8)$$

where $f(\mathbf{B}_{\text{PS}})$ is the rotation complexity function; λ_{ij} and $b_{\text{PS}(ij)}$ denote the estimated and hypothesized factor loadings; and w_{ij} is a loadings-specific weight that equals 0 when $b_{\text{PS}(ij)}$ is freely estimated and equals 1 when $b_{\text{PS}(ij)}$ equals 0. Note that this method only minimizes the squared discrepancies between the estimated and fixed (at zero) target loadings.

In a comprehensive review of factor rotation methods, Browne (2001) suggested that the partially specified target rotation could be improved if the rotation was applied iteratively "until the investigator is satisfied with the outcome" (p. 125). This suggestion was taken up by Moore et al. (2015), who proposed an objective convergence criterion when updating the partially specified target matrix. Specifically, during each iteration, a new (par-

tially specified) target matrix is generated until the pattern of zero loadings in the updated target of the n^{th} iteration is equivalent to that in the $(n - 1)^{\text{th}}$ iteration. The authors called their method an "iterated target rotation" (ITR; Moore et al., 2015, p. 149) and demonstrated the effectiveness of the ITR in recovering complex factor solutions. In later work, Abad et al. (2017) applied the ITR algorithm to bifactor models and called this application an iterated Schmid-Leiman (SLi) target rotation.

Jennrich and Bentler's Orthogonal and Oblique Bifactor Rotation Methods

Taking an approach that is more in line with traditional rotation algorithms, Jennrich and Bentler (2011, 2012, 2013)⁸ proposed algorithms for rotating factors to orthogonal and oblique EBF orientations. Both of their algorithms rely on the gradient projection algorithm (GPA; Jennrich, 2002). The first algorithm minimizes the orthogonal quartimin (Carroll, 1953, 1957) criterion whereas the second algorithm minimizes the oblique geomin (Yates, 1987; see also Browne, 2001) criterion for locating a simple structure solution for the k group factors of a $k + 1$ bifactor model.

In what follows we use the labels bifactor quartimin⁹ (BQ) and bifactor geomin (BG) to denote orthogonal or oblique exploratory bifactor rotations, respectively. In the orthogonal case, an orthogonal transformation matrix is found to produce a bifactor solution wherein all $k + 1$ factors are mutually orthogonal. To find this solution, Jennrich and Bentler (2011, 2013) minimize the quartimin (Carroll, 1953, 1957) criterion in the k group factors (note that in the Jennrich and Bentler approach, the general factor loadings do not influence the quartimin criterion) such that,

$$\text{qmin}(\mathbf{\Lambda}) = \sum_{i=1}^p \sum_{r=2}^{k+1} \sum_{s=r+1}^{k+1} \lambda_{ir}^2 \lambda_{is}^2, \quad (9)$$

where $\text{qmin}(\mathbf{\Lambda})$ is the quartimin criterion (see Jennrich & Bentler, 2011) to be minimized over the k group factors and λ_{ir} (λ_{is}) denotes an element in the i^{th} row and r^{th} (s^{th}) column of the $p \times (k + 1)$ matrix of factor loadings ($\mathbf{\Lambda}$). In short, the quartimin criterion seeks the factor solution that minimizes the row complexity of the structure matrix such that the discrepancy function will be minimized when the number of zero or near-zero row-wise elements is maximized (Carroll, 1953, 1957).

In the oblique geomin case, group factors are allowed to correlate with one another but not with the general factor. In Jennrich and Bentler's algorithm, a simple structure pattern in the group factor loadings is found by minimizing a modified (see Browne, 2001) version of the geomin criterion (Yates, 1987), such that

⁸ Note that Jennrich and Bentler (2013) published an errata for their 2011 article.

⁹ Jennrich and Bentler (2012) call their method bi-quartimin because it applies the quartimin criterion to the bifactor case. However, this method should not to be confused with Carroll's distinct bi-quartimin (Carroll, 1957) factor rotation criterion.

$$\text{gmin}(\Lambda) = \sum_{i=1}^p \left[\prod_{r=2}^{k+1} (\lambda_{ir}^2 + \epsilon) \right]^{\frac{1}{k}}, \quad (10)$$

where $\text{gmin}(\Lambda)$ denotes the geomin complexity function, λ_{ir} is an element in the i^{th} row and r^{th} column of the $p \times (k+1)$ bifactor pattern matrix, and ϵ is a small value (typically set to .01, see Hattori, Zhang, & Preacher, 2017).

Although it is not apparent in Equations (9) and (10), the general factor of the BQ and BG exploratory bifactor solutions is implicitly rotated during the projection step of the gradient projection algorithm (Mansolf & Reise, 2016). An unfortunate consequence “of the implicit rotation of the general factor is that the [Jennrich and Bentler] rotations are prone to local minima problems” (Mansolf & Reise, 2016, p. 707; see also Hattori et al., 2017). To demonstrate this point, Mansolf and Reise (2016) recovered two known bifactor loadings matrices, each exhibiting an independent cluster structure (McDonald, 1999, p. 174). To recover these bifactor models, they performed BQ and BG rotations from 1,000 random starting configurations. In the first model, Mansolf and Reise (2016) recovered five unique solutions for the BQ (bifactor quartimin) rotations and 42 unique solutions for the BG (bifactor geomin) rotations. In the second model, they recovered only one unique solution for the BQ rotation and 28 solutions for the BG rotations. Considered together, these results suggest that Jennrich and Bentler’s (2012, p. 444) recommendation to run their rotation methods from 10 random starting configurations is not sufficient to guarantee that a researcher will locate the global minimum.

Study 1: A Monte Carlo Study of Bifactor Recovery Using Seven Methods for Exploratory Bifactor Analysis

Method

Our Monte Carlo study was designed to compare the relative performance of seven exploratory bifactor rotation methods: (a) the SL (the Schmid-Leiman procedure; Schmid & Leiman, 1957); (b) the DSL (the Direct Schmid-Leiman method; Waller, 2018a); (c) the DBF (the Direct Bifactor method; Waller, 2018a); (d) the SLt (the partially specified target bifactor method described by Reise et al., 2010); (e) the SLi (the iterated partially specified target method proposed by Abad et al., 2017); (f) the BQ (the orthogonal, full-rank bifactor quartimin rotation developed by Jennrich & Bentler, 2011, 2013); and (g) the BG (the oblique, full-rank bifactor geomin rotation developed by Jennrich & Bentler, 2012). Our overarching goals for this study were threefold. First, we wished to compare the relative performance of the recently proposed DSL and DBF (Waller, 2018a) methods against the more thoroughly studied exploratory bifactor methods that were evaluated in the Abad et al. (2017) study. Second, we wished to test the ability of all of the aforementioned bifactor methods to recover known bifactor solutions across both hierarchical and nonhierarchical bifactor models. Previous work in this area (Abad et al., 2017) focused exclusively on the recovery of nonhierarchical models despite the popularity of hierarchical bifactor structures in the applied literature (e.g., Dombrowski, McGill, & Canivez, 2017; Serra, Spoto, Ghisi, & Vidotto, 2017; Swanson, 2017). Finally, we wished to extend the simulation design of Abad et al.

(2017) in two additional important respects. Namely, relative to the Abad et al. (2017) study, our design includes population bifactor models with a larger range of group factors, {3, 5, 10}, and a larger range of indicators per group factor {4, 6, 8}.

Simulation design. Our simulations included population bifactor models with the following characteristics: (a) number of variables per group factor, VAR.GrpFac \in {4, 6, 8}; (b) number of group factors, NUM.GrpFac \in {3, 5, 10}; (c) magnitude of group factor cross-loadings, CROSS.GrpFac \in {0, .40}; (d) magnitude of general factor loadings, LOAD.GenFac \in {.30–.50, .40–.60, .50–.70}; (e) magnitude of group factor loadings, LOAD.GrpFac \in {.30–.50, .40–.60, .50–.70}; (f) population model type {hierarchical, nonhierarchical}; and (g) sample size, $N \in$ {200, 500, 2,000}. In aggregate, these seven design variables produce a $3 \times 3 \times 2 \times 3 \times 3 \times 2 \times 3$ fully crossed, factorial design with 972 conditions.

Our design deviated from that in Abad et al. (2017) in two important aspects. First, our population bifactor models did not include items that were pure markers of the general factor only (i.e., items that load on the general factor but not any group factor). Pure general factor markers were excluded from our simulations because items that load on only a general factor will not be found in hierarchical bifactor models (due to the loadings constraints that are implicit in Equation 3). Thus, we omitted this condition to retain our fully crossed factorial design. Second, we broadened the ranges of two design factors that were included in the Abad et al. (2017) study. Specifically, whereas Abad et al. (2017) generated models with four, five, or six indicators per group factor, our models included four, six, or eight indicators per group factor. Moreover, whereas Abad et al. (2017) generated models with four, five, or six group factors, our models included three, five, or 10 group factors. We expanded the range of these variables to better align our design with published applications of bifactor models. For instance, recent applications of the bifactor model have examined models with six or more group factors (e.g., Cucina & Howardson, 2017; Litalien et al., 2017) and six or more indicators per group factor (e.g., Cucina & Howardson, 2017).

Data generation. For each condition of our design, data were generated as follows. First, a population bifactor structure was generated from the aforementioned design levels. All general factor loadings were initially drawn in equal increments within the lower and upper bounds of the three previously described levels (i.e., Gen $\lambda \in$ {.30–.50, .40–.60, .50–.70}) and then randomly permuted. The length of the general factor equaled the number of indicators per group factor multiplied by the number of group factors in the model. All group factor loadings were likewise drawn in equal increments (between specified bounds) and retained in ascending order. In models that included cross-loadings, the placements of the cross-loadings were chosen as follows. For each group factor, the last factor indicator had a secondary loading of .40 on the subsequent group factor (or on the first group factor when the indicator was a primary marker of the last group factor). To reduce the potential for Heywood cases (i.e., item communalities > 1.0), all general factor and primary-group factor loadings for items with secondary factor loadings were rescaled to maintain the pre-cross-loading communality while retaining a cross-loading of .40.

For each model in our design, we generated a population correlation matrix, such that,

$$\mathbf{R}_{\text{pop}} = \mathbf{\Lambda}\mathbf{\Lambda}^T + \mathbf{\Theta}^2, \quad (11)$$

where \mathbf{R}_{pop} denotes $p \times p$ model-implied correlation matrix, $\mathbf{\Lambda}$ denotes the $p \times (k + 1)$ population, nonhierarchical bifactor-loadings matrix, and $\mathbf{\Theta}^2$ is a $p \times p$ diagonal matrix of uniqueness factor variances so chosen to constrain the diagonal elements of \mathbf{R}_{pop} to one. An additional step was required to convert the nonhierarchical bifactor models into hierarchical bifactor models. Specifically, designating a nonhierarchical bifactor structure matrix ($\mathbf{\Lambda}$) as a target matrix, we used the DSL procedure to find a hierarchical structure with the smallest Euclidean distance to the target (nonhierarchical) structure. To quantify the degree of divergence between the nonhierarchical model and its associated hierarchical structure, we compared the difference between the nonhierarchical and hierarchical bifactor structures for each condition that influences the population structure (i.e., 162 conditions). In this small simulation study, the median divergence between the full-rank (nonhierarchical) and rank-deficient (hierarchical) models, as measured by RMSE, was .02 ($M = .03$, $\max = .07$; see the [online supplemental materials](#) for full results). Moreover, the median congruence coefficients between all model pairs was 1.00 ($M = 1.00$; all results rounded to two decimal places). [Table 1](#) reports the closest hierarchical bifactor model to a representative nonhierarchical model from our simulations. In this example, the RMSE for the two solutions was .02 (average congruence coefficient = .998). Notice in [Table 1](#) that the overall structures of the two solutions are highly similar.

For each pair of nonhierarchical and hierarchical (population) bifactor loadings matrices, we used a procedure described by [Browne \(1968\)](#); see also, [Kshirsagar, 1959](#)) to generate 50 sample correlation matrices for each of the three sample sizes (i.e., 200, 500, or 2,000) included in our design.¹⁰ For each sample correlation matrix, we estimated an exploratory least squares bifactor analysis model using seven methods: (a) the SL (the Schmid-Leiman procedure; [Schmid & Leiman, 1957](#)); (b) the exploratory DSL (the Direct Schmid-Leiman method; [Waller, 2018a](#)); (c) the exploratory DBF (the Direct Bifactor method; [Waller, 2018a](#)); (d) the SLt (the partially specified

target bifactor method described by [Reise et al., 2010](#)); (e) the SLi (the iterated partially specified target method proposed by [Abad et al., 2017](#)) using the SL and DSL final solutions as different starting configurations; (f) the BQ (the orthogonal, full-rank bifactor quartimin rotation developed by [Jennrich & Bentler, 2011, 2013](#)); and (g) the BG (the oblique, full-rank bifactor geomin rotation developed by [Jennrich & Bentler, 2012](#)).

Criterion for model recovery. After fitting bifactor models to the sample data sets, we quantified model recovery by computing a root mean squared error (RMSE) between the population and estimated bifactor loadings matrices. To make our findings comparable to those reported in the [Abad et al.'s \(2017\)](#) study, we also quantified model recovery by computing average congruence coefficients (CC; [Tucker, 1951](#)) between the population and estimated bifactor loadings matrices. The congruence coefficient results were highly consistent with the RMSE findings, and thus we report these additional (CC) recovery findings in the electronic [online supplemental materials](#).¹¹

Before computing an overall RMSE for each pair of population and estimated bifactor models, it was necessary to align the factors from the estimated solution with those from the population model. In the present study, we accomplished this alignment by locating the factor permutation and orientation (of the estimated factors) that minimized the RMSE between the two structure matrices. Our method guaranteed that the chosen alignment represented the global minimum RMSE value.

All simulations were conducted in R ([R Core Team, 2018](#)) with functions written by us (and in some cases modified from code reported by [Abad et al., 2017](#)) for the SL, Direct SL, Direct Bifactor, SLt, and SLi algorithms, with the bifactor quartimin and bifactor geomin functions found in the R `GPArotation` library ([Bernaards & Jennrich, 2005](#)). All factor analyses were conducted using an unweighted least squares discrepancy function and all Schmid-Leiman and exploratory DSL and DBF analyses were performed using geomin ([Yates, 1987](#)) rotations.¹²

Before leaving this section, it merits comment that recent discussions of factor rotations (e.g., [Browne, 2001](#); [Hattori et](#)

Table 1
The Closest Hierarchical Bifactor Structure to a Given Nonhierarchical Bifactor Structure

Item	Nonhierarchical				Hierarchical			
	1	2	3	4	1	2	3	4
1	.61	.50	.00	.00	.58	.53	.04	.03
2	.50	.57	.00	.00	.51	.56	-.01	-.01
3	.54	.63	.00	.00	.55	.62	-.02	-.02
4	.65	.70	.00	.00	.65	.69	-.01	-.01
5	.55	.00	.50	.00	.55	.00	.50	.00
6	.70	.00	.57	.00	.68	.02	.59	.02
7	.66	.00	.63	.00	.67	-.01	.63	-.01
8	.68	.00	.70	.00	.70	-.02	.68	-.02
9	.63	.00	.00	.50	.59	.04	.04	.53
10	.57	.00	.00	.57	.57	.01	.01	.58
11	.59	.00	.00	.63	.59	.00	.00	.63
12	.52	.00	.00	.70	.55	-.03	-.04	.66

Note. Loadings $> |.10|$ are printed in boldface.

¹⁰ The correlation matrices were generated (using the `corSample` function from the `fungible` R library, [Waller, 2018b](#)) by drawing covariance matrices from an appropriate Wishart distribution and then scaling the covariances into correlations. See [Kshirsagar \(1959\)](#) for details.

¹¹ In this article we quantified model recovery via the RMSE rather than with congruence coefficients because of well-known problems with the latter measure of matrix similarity. For example, [Pinneau and Newhouse \(1964\)](#) report two single-factor models in which the first model has loadings of (.01, .03, .02, .04) and the second model has loadings of (.91, .93, .92, .94). As noted by these authors, the congruence coefficient between these factors is .92. According to some researchers, congruence coefficients larger than .90 are considered indicative (e.g., [Lorenzo-Seva & ten Berge, 2006](#)) of fair similarity between the two solutions. Nevertheless, as clearly shown in this example, these factor loadings are highly discrepant and have a mean difference of .90.

¹² All geomin rotations used an ϵ (referred to as δ in the `GPArotation` library) tuning parameter of .01 as suggested by [Hattori et al. \(2017\)](#). We chose geomin for the current study to more easily compare our results with those of [Abad et al. \(2017\)](#) who also used geomin rotation for their Schmid-Leiman analyses.

al., 2017; Mansolf & Reise, 2016; Rozeboom, 1992) have highlighted that many rotation algorithms can converge on local minima. To address this issue, we conducted all factor rotations from 100 (rather than 10 as recommended by Jennrich & Bentler, 2012) random starting configurations and chose the solution with the lowest criterion value. Lastly, the Jennrich and Bentler (2012) rotations were computed with Kaiser normalization.

Results

To summarize our main results, we computed an analysis of variance (ANOVA) for each method using the RMSE values as the dependent variable and the aforementioned design factors separately for the nonhierarchical and hierarchical bifactor models. For each ANOVA, we computed main effects, two-way interactions, and three-way interactions of the six manipulated variables. Each cell of our design was replicated 50 times.

A summary of the ANOVA results is displayed Table 2. For brevity, the table only reports model terms with effect sizes (η^2) greater than or equal to .02 (i.e., explaining at least 2% of the variance of the RMSE values; Cohen, 1973, 1988) in at least one of the exploratory bifactor methods. Note that in Table 2, we report results for SLi from an SL starting configuration but omit results for the SLi using a DSL starting configuration. In our simulations, the performance of the two SLi methods was nearly identical.

As shown in Table 2, the pattern of effect sizes across the two model types—that is, the nonhierarchical and hierarchical population structures—was highly similar. Notice that for all methods, except the Jennrich and Bentler rotations, sample size produced the largest effects ($.32 \leq \eta^2 \leq .71$) and for many methods, number of group factors and group factor loading strength produced large effects. For the BQ and BG methods,

however, number of group factors accounted for the lion's share of the dependent variable variances in both the nonhierarchical and hierarchical bifactor models. Notice also that many interaction terms accounted for a modest portion of variance with the strongest interaction occurring for the DSL method when there are differences between the loadings on the general and group factors. This effect size likely reflects the fact that the DSL target matrix is composed of ones and zeros regardless of the loading variation in the initial factor solution that is used to approximate \mathbf{B}_{pop} in (7). Finally, notice that in the nonhierarchical models, the presence of cross-loadings accounted for significant proportions of variance in all methods except the SLt and SLi procedures. These latter methods appear not to be affected by this model characteristic and this general finding is consistent with that reported by Abad et al. (2017). Nevertheless, a different picture emerges when we consider the results from hierarchical bifactor models. In these cases, only the Jennrich and Bentler methods (i.e., BQ and BG) were influenced by the presence of cross-loadings. To further understand these findings, we generated a number of visual displays to elucidate the interactions among several of our design factors.

Visual displays of bifactor model recovery. Although the results from ANOVAs can be informative, indices of explained variance (e.g., η^2) do not directly address the main question that motivated our study. To wit: How well do the various EBF methods stack up when recovering nonhierarchical and hierarchical bifactor models? To answer this question, we generated a series of conditioning plots (coplots; Becker, Cleveland, & Shyu, 1996) to compare model recovery for five of the EBF methods that were included in our study. Notice in these figures that we do not report results for the Schmid-Leiman method because (a) in general, for the simulations in Study 1, the SL

Table 2
 η^2 Effect Sizes for All Design Variables and Bifactor Methods

Model	Independent variable	SL	DSL	DBF	SLt	SLi	BQ	BG
Nonhierarchical	<i>N</i>	.49	.39	.51	.55	.48	.12	.08
	VAR.GrpFac	.02	.02	.01	.02	.02	.02	.04
	NUM.GrpFac	.10	.14	.14	.09	.11	.38	.51
	CROSS.GrpFac	.12	.12	.02	.01	.01	.10	.12
	LOAD.GrpFac	.07	.07	.07	.05	.06	.00	.00
	LOAD.GenFac	.02	.02	.04	.05	.05	.00	.01
	<i>N</i> × LOAD.GrpFac	.02	.02	.02	.01	.01	.00	.00
	NUM.GrpFac × CROSS.GrpFac	.03	.03	.01	.00	.00	.04	.04
	LOAD.GrpFac × LOAD.GenFac	.00	.04	.03	.00	.00	.00	.00
	<i>N</i>	.71	.57	.48	.43	.32	.05	.06
Hierarchical	VAR.GrpFac	.01	.01	.02	.03	.03	.01	.03
	NUM.GrpFac	.00	.02	.20	.20	.24	.43	.55
	CROSS.GrpFac	.00	.01	.00	.00	.01	.08	.06
	LOAD.GrpFac	.08	.10	.08	.04	.05	.01	.00
	LOAD.GenFac	.03	.03	.03	.04	.03	.00	.01
	<i>N</i> × LOAD.GrpFac	.02	.03	.01	.01	.01	.01	.01
	NUM.GrpFac × CROSS.GrpFac	.00	.00	.00	.00	.01	.04	.02
	NUM.GrpFac × LOAD.GrpFac	.01	.00	.00	.00	.00	.03	.02
	LOAD.GrpFac × LOAD.GenFac	.00	.08	.03	.00	.00	.00	.00
	NUM.GrpFac × LOAD.GrpFac × LOAD.GenFac	.00	.04	.01	.00	.00	.00	.00

Note. Each column within the model is a single ANOVA for that method. All values greater than or equal to .02 are printed in boldface. *N* = sample size; VAR.GrpFac = number of variables per group factor; NUM.GrpFac = number of group factors; CROSS.GrpFac = presence of cross-loadings; LOAD.GrpFac = magnitude of the group factor loadings; LOAD.GenFac = magnitude of the general factor loadings; SL = Schmid-Leiman; DSL = Direct Schmid-Leiman; DBF = Direct Bifactor; SLt = Schmid-Leiman target; SLi = Schmid-Leiman iterated; BQ = bifactor quartimin; BG = bifactor gomin.

and DSL methods produced relatively similar findings (with some notable exceptions that are summarized in Tables 3 and 7, and Figures 5, 6, 15, and 16 in our online supplemental material); and (b) later, we provide a more detailed comparison of the SL and DSL methods in a second simulation study. Note that we also do not report the results for the SLt because of the high similarity (but lower performance of the SLt) of these findings with the SLi.¹³ The initial coplots from these analyses are reported in Figure 1. This figure includes boxplots of the RMSE recovery values for the (a) DSL, (b) DBF, (c) SLi, (d), BQ, and (e) BG methods when the data have been stratified by group-factor number and sample size. The results for the nonhierarchical and hierarchical models have been separated to facilitate interpretation of the model interaction terms.

When scanning the findings in Figure 1, notice that the coplots include horizontal reference lines at RMSE values of .05 and .15. These RMSE values should not be interpreted as benchmarks for model recovery but only as visual guides to ease cross-panel comparisons. Let us begin by considering the findings for the nonhierarchical bifactor models that are summarized in Figure 1A. To generate this figure we conditioned the results on number of group factors (labeled “# GrpFac” in this figure) and sample size (N). We chose to focus on these conditioning variables because they were associated with large effect sizes in our aforementioned ANOVA results (see Table 2). Scanning across the panels in Figure 1A, two findings stand out. First, relative to the other methods in our simulation, DSL provides superior model recovery in all but the largest sample sizes. This seems to be another example in which unit weights perform well in small to moderately sized samples (Wainer, 1976, 1978). The second observation that merits comment concerns the relatively poor performance of the two Jennrich and Bentler (2011, 2012) methods (i.e., bifactor quartimin and bifactor geomin) under most conditions (cf. Abad et al., 2017). Indeed, the only conditions in which BQ and BG perform reasonably well is when both the sample size and number of group factors are unusually large.

Zooming in on Figure 1A, we see that SLi performs slightly better than DSL, on average, in the largest sample size ($N = 2,000$), but it performs less well than DSL in terms of a minimax criterion in which one considers the maximum RMSE values. Thus, even in very large samples it may be preferable to estimate nonhierarchical bifactor models with the rank-deficient DSL procedure. On first glance, this is a surprising result. However, as we noted earlier, rank-deficient DSL models can often closely approximate nonhierarchical bifactor models of a higher rank. Thus, when considered in aggregate, our findings strongly support the advice that researchers should estimate the parameters of full-rank, nonhierarchical bifactor models with the DSL procedure in small to moderately sized samples.

Let us now consider the relative performance of the five estimation methods for recovering the parameters of hierarchical bifactor models conditioning on number of group factors and sample size. Figure 1B displays boxplots of the RMSE recovery values for the five EBF methods that are included in Figure 1A. Notice that in all conditions DSL outperformed the other methods when performance was gauged by either the typical (i.e., average or median RMSE) or the minimax criteria. On the one hand, this result was expected given that DSL is the only rank-deficient method in this set of EBF methods. On the other hand, our

previous results summarized in Figure 1A cogently demonstrate that expectations and outcomes need not coincide. In summary, when recovering hierarchical bifactor models, our results indicate that DSL exhibits excellent model recovery, especially in models with few group factors. In models with a relatively larger number of group factors (e.g., # Grp Fac = 10) that are estimated in moderate to large samples, all methods in these simulations performed well on average. However, even in the largest sample sizes, the full-rank methods performed exceedingly bad in some cases.

Figure 2 displays model-recovery coplots for the nonhierarchical and hierarchical bifactor models when conditioning on group-factor loading range (labeled “Grp λ ”) and sample size (N). Scanning both plots in Figure 2, we see similar trends to those in Figure 1. Specifically, across most conditions, DSL produced the best model recovery of the five EBF methods. This was true in terms of both the average RMSE and the maximum RMSE values. In some conditions, SLi performed slightly better than the DSL in terms of average performance, but it never outperformed the DSL in terms of the minimax criterion. Interestingly, SLi performance does not seem to improve as a function of communality strength (cf. MacCallum, Widaman, Preacher, & Hong, 2001; MacCallum, Widaman, Zhang, & Hong, 1999). Finally, consistent with our earlier findings, the results summarized in Figure 2 show that the two Jennrich and Bentler algorithms (BQ and BG) rarely performed well in these simulations.

Earlier we noted that in Study 1, SL and DSL produced highly similar results and thus, for the sake of brevity, we omitted the SL findings from our discussion. We also noted in our description of the DSL and DBF that there are several variants of these methods and that the variants included in Study 1 used an empirically generated target matrix to approximate the population bifactor model (in Equation 7). We now take a closer look at these noniterative EBF methods and we compare their relative performance to the original Schmid-Leiman method.

Study 2: A Closer Look at Three Noniterative Methods for Exploratory Bifactor Analysis

Readers may recall that in Study 1 we computed empirically generated target matrices for the DSL by (a) dichotomizing an obliquely rotated factor matrix from a prior exploratory factor analysis and then (b) appending a vector of ones to this dichotomized loadings matrix (see Two Direct Schmid-Leiman Methods section). These admittedly crude target matrices consisted of (signed) ones and zeros. When investigators have more knowledge about the pattern of population bifactor loadings, it would be desirable to incorporate this information when constructing a target matrix. Following up on this idea, in the present study we compared three methods for constructing DSL target matrices that use various amounts of prior information concerning the population loadings structure. The first method is used in fully exploratory studies in which an investigator has minimal knowledge about the pattern (and magnitudes) of the bifactor loadings. This method was used in all of the DSL analyses that were carried out in Study 1. A second method for constructing DSL target matrices takes advantage of investigator knowledge by allowing the inves-

¹³ The online supplemental materials summarize the results for all methods.

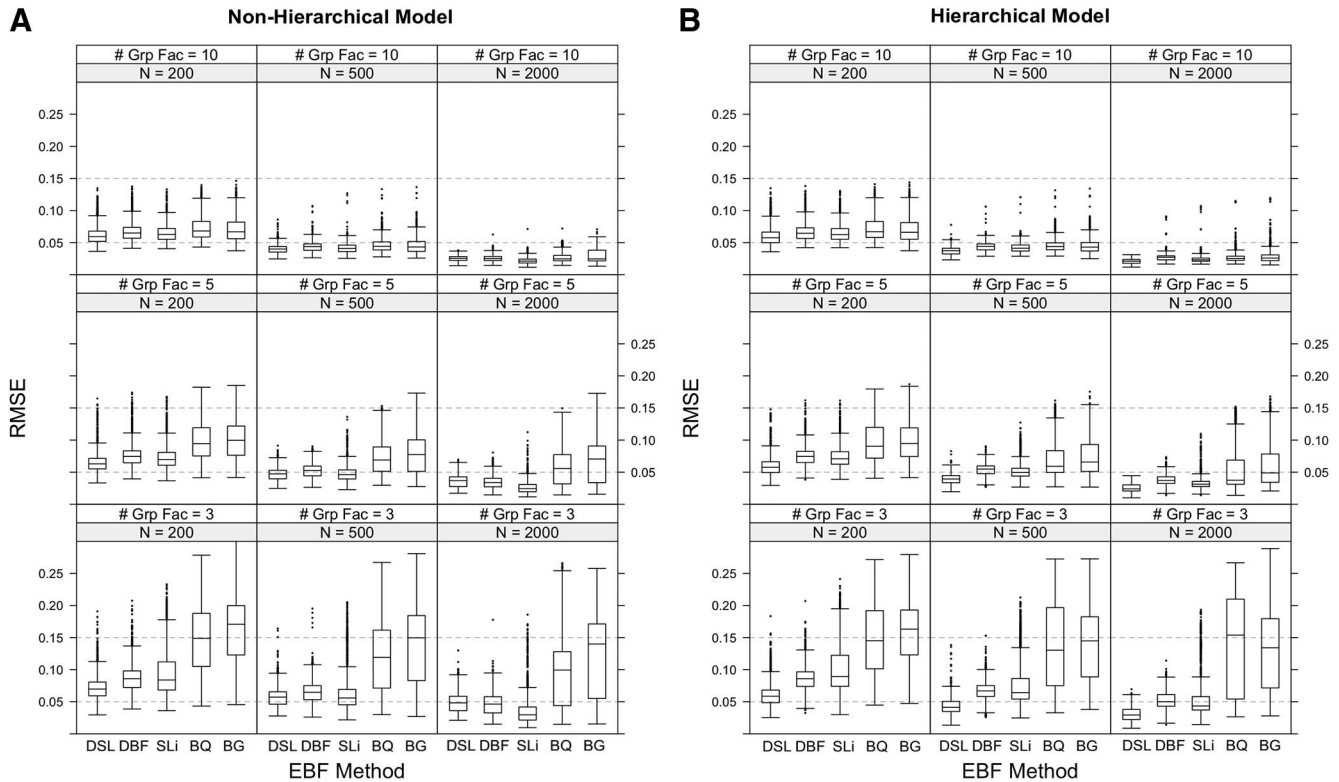


Figure 1. Bifactor model recovery coplots for sample size and number of group factor. # Grp Fac = number of group factors; N = sample size; RMSE = root mean squared residual; DSL = Direct Schmid-Leiman; DBF = Direct Bifactor; SLi = Schmid-Leiman iterated target rotation; BQ = bifactor quartimin; BG = bifactor geomin; EBF = exploratory bifactor.

tigator to specify a pattern of (signed) ones and zeros when constructing a target matrix. The third method for constructing a DSL target matrix uses information from prior analyses (or privileged knowledge) of the content domain. Specifically, in this method, the bifactor loadings matrix from a previous study, or the known population loadings in a simulation study, are used as the target matrix. When population loadings are known in advance, this method for constructing a DSL target matrix provides a useful benchmark for comparing the relative performances of other methods for conducting exploratory bifactor analysis. In prior work, Waller (2018a) showed that no other rank-deficient EBF method can outperform the DSL (in terms of minimizing RMSE) with a population target matrix. Moreover, as shown in Study 1, this third method of creating DSL target matrices has useful applications in Monte Carlo work because it can be used to generate the closest hierarchical bifactor model to a known nonhierarchical bifactor model.

In the current study we compared six cognate bifactor estimation methods in their ability to recover known bifactor structures: (a) the Schmid-Leiman method (SL); (b) the exploratory Direct Schmid-Leiman (DSL) method with an empirically generated target matrix; (c) a quasiconfirmatory DSL method (DSL_±) in which known population loadings are dichotomized to create a signed target matrix; (d) a quasiconfirmatory DSL method (DSL_p) in which population loadings are used as the target matrix; (e) a quasiconfirmatory DBF method (DBF_±) in which population load-

ings are dichotomized to create a signed target matrix; and (f) a quasiconfirmatory DBF method (DBF_p) in which known population loadings are used as the target matrix. We compared these methods using a design that was similar to that used in Study 1. In Study 2, we also generated sample data sets for each of the 972 factor conditions that were described in Study 1 but, in contrast to the earlier study, in the present study we generated 500 data sets, rather than 50, for each condition. The additional replications were included to allow us to more reliably differentiate between the several variants of the Direct Schmid-Leiman method.

Figure 3 summarizes the ability of each method to recover the nonhierarchical bifactor models that were included in these simulations. Notice in this figure that the Schmid-Leiman and exploratory DSL methods produced similar results on average but there were also differences that merit discussion. Specifically, when sample sizes were small to moderate and group factor loadings were high, relative to the SL procedure, DSL produced less extreme RMSE outliers. However, in the largest sample sizes (summarized in the right-most column of Figure 3), SL performed slightly better than DSL under all levels of group-factor loadings size. These results suggest that, unless sample sizes are exceedingly large (e.g., $N \geq 2,000$), the exploratory DSL offers a slight advantage over the Schmid-Leiman method for recovering nonhierarchical factor structures.

The findings in Figure 3 also show that, when population loadings are known, using an optimal target matrix (i.e., the binary

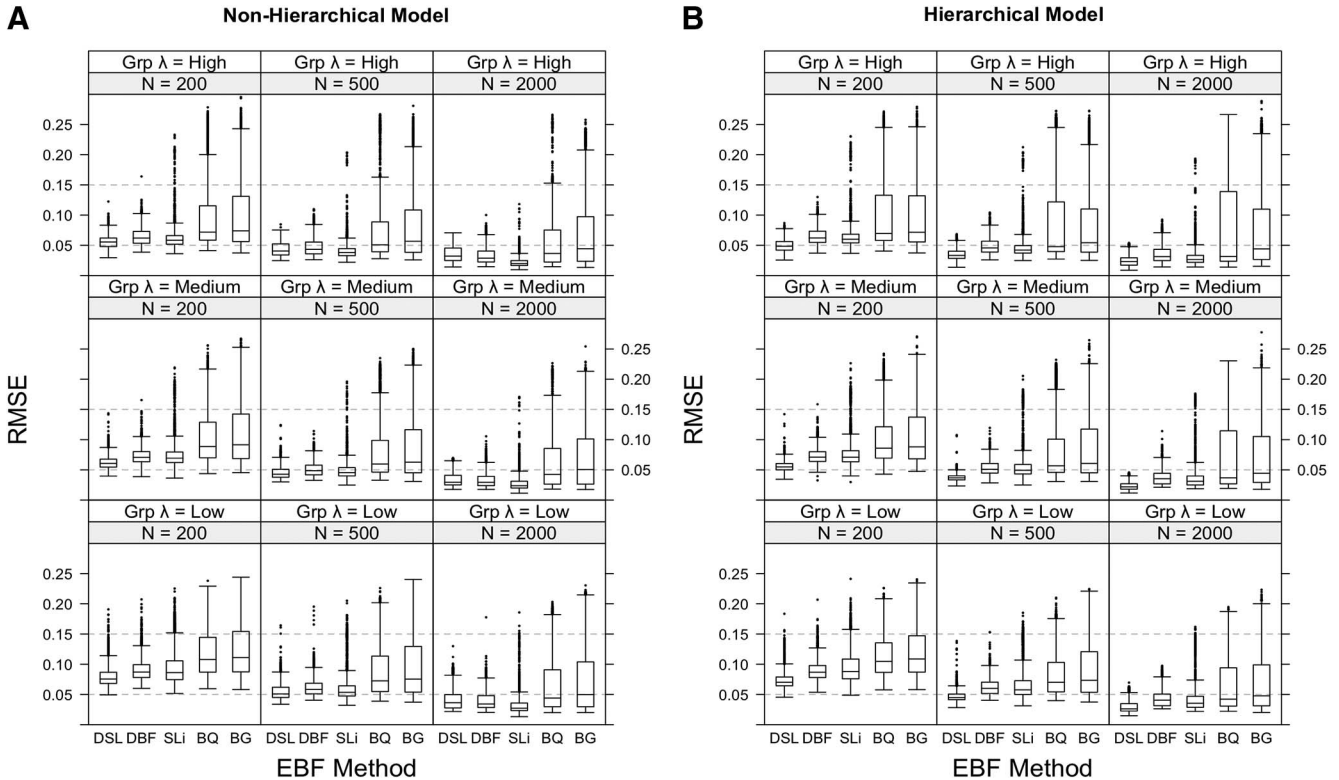


Figure 2. Bifactor model recovery coplots for sample size and group factor loading strength. Grp λ = magnitude of the group-factor loading range; Low = $.30 > \lambda > .50$; Medium = $.40 > \lambda > .60$; High = $.50 > \lambda > .70$; N = sample size; RMSE = root mean squared residual; DSL = Direct Schmid-Leiman; DBF = Direct Bifactor; SLi = Schmid-Leiman iterated target rotation; BQ = bifactor quartimin; BG = bifactor geomin; EBF = exploratory bifactor.

or analogue population loadings matrix) does not generally improve the average performance of these methods. Rather, the use of an optimal target matrix tends to reduce the number of outlying RMSE values for the DSL. Thus, one of the most encouraging findings from these analyses is that there is only a small loss in model recovery when using the fully exploratory DSL rather than the quasiconfirmatory variants of this method. A second notable finding that can be gleaned from Figure 3 concerns the relative performance of the DSL_p and DBF_p. Although intuition might suggest that the full-rank DBF_p should outperform the rank-deficient DSL_p when recovering nonhierarchical bifactor models, intuition is apparently wrong in this case.

Moving on to the hierarchical bifactor models, Figure 4 summarizes the model recovery values of the three EBF methods (and their variants) for these simulations. Because both the SL and DSL methods were specifically designed to recover hierarchical bifactor structures, we expected that both methods would perform well in these simulations. The results shown in Figure 4 confirmed our expectation. Note that under most conditions, the associated SL and DSL RMSE values were both comparable and small. Moreover, consistent with earlier results, in all but the largest sample sizes the DSL showed better minimax performance. Under the largest sample sizes, SL performance was slightly superior to that of DSL, but the recovery differential was trivial in these cases. Also consistent with our previous results, we found that providing

DSL with an increasingly optimal target matrix had little effect on typical method performance but a more noticeable effect on the number and range of the outlying RMSE values. One of the more noteworthy results of these EBF comparisons concerns the relatively poor performance of the full-rank DBF methods. Regarding these results, notice that even when the DBF is given a known population target matrix, it still performed worse than the exploratory DSL method in all models with moderate to high group factor loadings.

General Discussion

The last decade has witnessed a resurgence of interest in exploratory bifactor analysis (EBF) models and the concomitant development of new methods to estimate these models (Abad et al., 2017; Jennrich & Bentler, 2011, 2012, 2013; Reise, 2012; Reise et al., 2010; Waller, 2018a). Until recently, researchers wishing to conduct an exploratory bifactor analysis had few analytic choices (e.g., Schmid & Leiman, 1957; Wherry, 1959). Today's researchers, however, have many more options due to new algorithms described by (a) Jennrich and Bentler (2011, 2012, 2013); (b) Reise et al. (2010); (c) Abad et al. (2017; see also Moore et al., 2015); and (d) Waller (2018a). Given this cornucopia of choices, investigators are now faced with a more difficult decision—namely, choosing an optimal method for exploratory bifac-

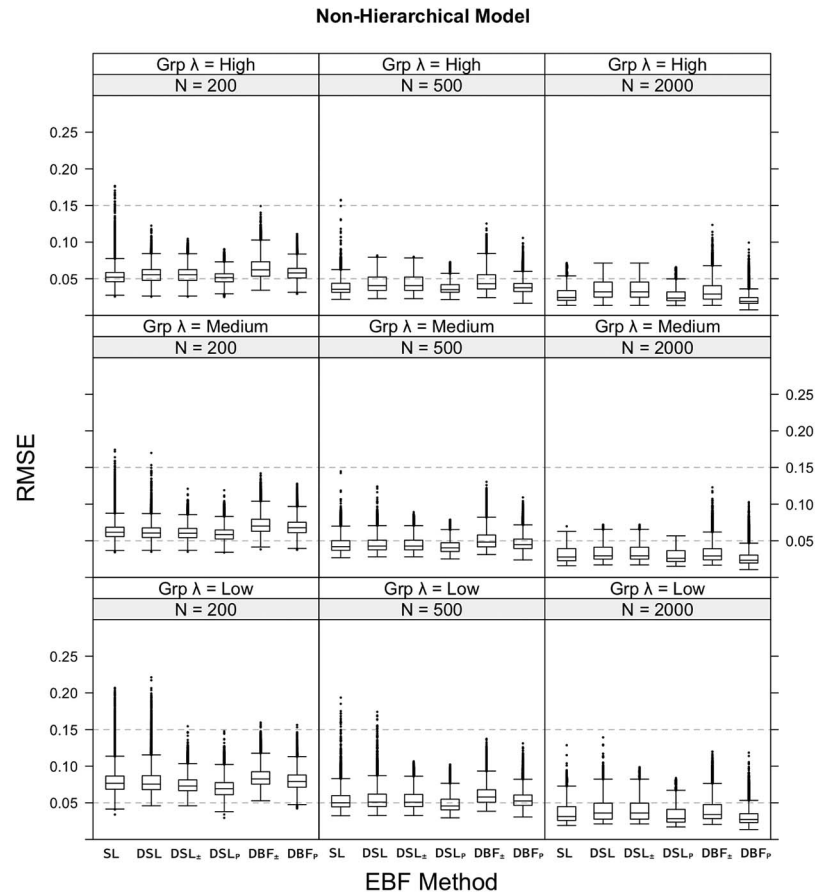


Figure 3. Nonhierarchical bifactor model recovery coplots for sample size and group factor loading strength in six EBF methods. Grp λ = magnitude of the group-factor loading range; Low = $.30 > \lambda > .50$; Medium = $.40 > \lambda > .60$; High = $.50 > \lambda > .70$; N = sample size; RMSE = root mean squared residual; SL = Schmid-Leiman; DSL = Direct Schmid-Leiman; DSL $_{\pm}$ = DSL with a population-based signed target matrix; DSL $_p$ = DSL with the population loadings as a target matrix; DBF $_{\pm}$ = Direct Bifactor with a population-based signed target matrix; DBF $_p$ = DBF with the population loadings as a target matrix.

tor analysis given the dearth of Monte Carlo studies to guide method choice.

To our knowledge, only one comprehensive Monte Carlo study of exploratory bifactor analysis has appeared in the literature. This study was conducted by Abad et al. (2017) who simulated data for 972 population bifactor models to compare the model recovery capabilities of (a) the Schmid-Leiman method (Schmid & Leiman, 1957); (b) the Schmid-Leiman target method (Reise et al., 2010); (c) the iterated Schmid-Leiman target method (Abad et al., 2017; Moore et al., 2015); and (d) the orthogonal and oblique Jennrich and Bentler (2011, 2012, 2013) bifactor rotation algorithms. According to these authors, the iterated Schmid-Leiman target method was “the most consistent and accurate of the bi-factor rotations considered” (Abad et al., 2017, p. 11).

Although the Abad et al. (2017) study significantly expanded our understanding of five bifactor analysis methods, their study had two notable limitations. First, their Monte Carlo design did not include hierarchical bifactor models. In our opinion, this is a limitation because hierarchical bifactor models are often used to describe the latent structure of psychological constructs (e.g.,

Brunner, Nagy, & Wilhelm, 2012; Carroll, 1993; Markon, Krueger, & Watson, 2005). Second, the Abad et al. (2017) study—through no fault of the authors—did not include two recently proposed methods (Waller, 2018a) for exploratory bifactor analysis. To address these issues, we conducted an expanded Monte Carlo study of exploratory bifactor analysis. Our study included over 500,000 data sets representing 972 population models and we conducted nearly three million bifactor analyses.

Our results can be easily summarized. Namely, across most conditions the Schmid-Leiman (Schmid & Leiman, 1957) and the direct Schmid-Leiman (Waller, 2018a) methods were best able to recover the true population factor loadings in both hierarchical and nonhierarchical bifactor models. Counterintuitively, although both methods were designed to fit (rank-deficient) hierarchical bifactor models, they generally outperformed the other methods (Abad et al., 2017; Jennrich & Bentler, 2011, 2012, 2013; Reise, 2012; Reise et al., 2010; Waller, 2018a) when fitting both hierarchical and nonhierarchical bifactor models. Overall, these methods (a) showed the best typical performance and (b) produced the fewest and least extreme outliers in their fit distributions.

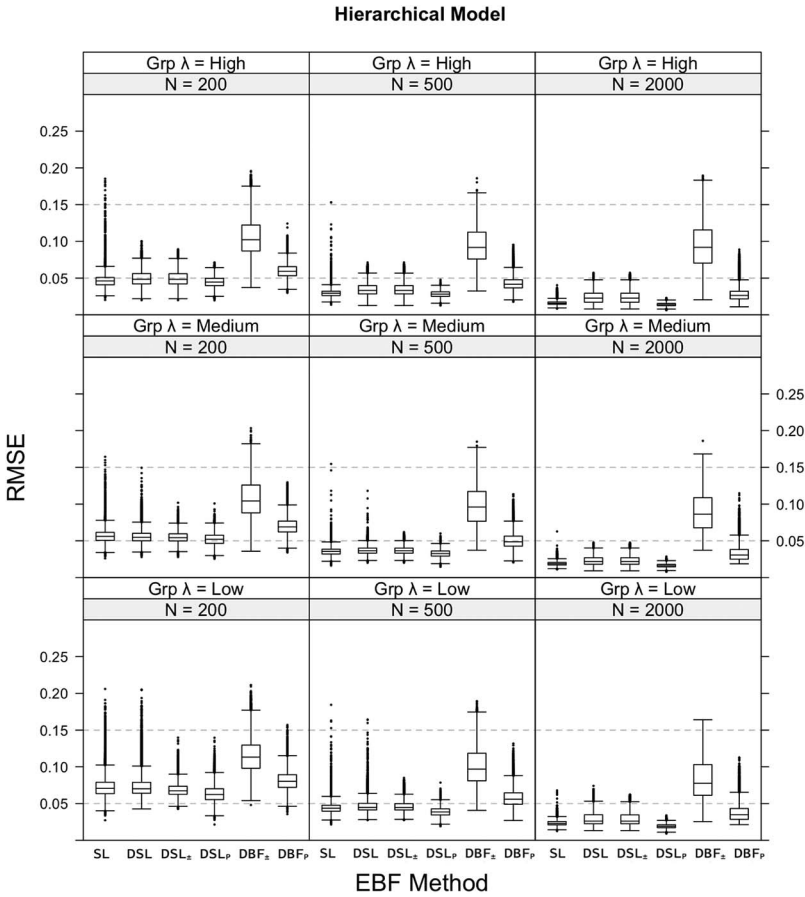


Figure 4. Hierarchical bifactor model recovery coplots for sample size and group factor loading strength in six EBF methods. Grp λ = magnitude of the group-factor loading range; Low = $.30 > \lambda > .50$; Medium = $.40 > \lambda > .60$; High = $.50 > \lambda > .70$; N = sample size; RMSE = root mean squared residual; SL = Schmid-Leiman; DSL = Direct Schmid-Leiman; DSL $_{\pm}$ = DSL with a population-based signed target matrix; DSL $_p$ = DSL with the population loadings as a target matrix; DBF $_{\pm}$ = Direct Bifactor with a population-based signed target matrix; DBF $_p$ = DBF with the population loadings as a target matrix.

Our results also suggest that when comparing the recovery capabilities of the Schmid-Leiman (Schmid & Leiman, 1957) and the Direct Schmid-Leiman (Waller, 2018a) methods there is no unequivocal winner per se. However, our findings indicate that in small to moderately sized samples (e.g., $N \leq 500$), researchers should choose the Direct Schmid-Leiman (Waller, 2018a) method to fit hierarchical or nonhierarchical exploratory bifactor models. Considering that many samples in applied psychology include fewer than 200 subjects (Carroll, 1993; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Shen et al., 2011), this recommendation seems well justified.

One notable limitation of our study is that—like others before us (cf. Abad et al., 2017; Mansolf & Reise, 2016)—we did not include nonbifactor population models in our Monte Carlo design. One reviewer of this article (William Revelle) astutely noted that an exploratory DSL analysis will suggest the presence of a hierarchical bifactor model when, in fact, the population model consists of orthogonal factors with an ideal, independent-cluster loadings pattern (i.e., when the factors are uniformly uncorrelated and the factor structure does not include cross-loadings). Under this

scenario, SL and DSL will yield markedly different solutions as the SL analysis will not produce a pseudo general factor. To address this limitation, in collaboration with Revelle, we are currently conducting a new Monte Carlo study that is specifically designed to identify the conditions under which false evidence for bifactor models will be produced by exploratory and confirmatory bifactor measurement procedures. Some preliminary findings from these simulations suggest that when an orthogonal factor structure includes cross-loadings, and/or when one is analyzing sample rather than population data (from an orthogonal factor model), then both the SL and DSL methods can produce false evidence for bifactor structures. This is a topic that surely warrants further exploration.

References

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

- Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, 3, 121–136. <http://dx.doi.org/10.3390/jintelligence3040121>
- Becker, R. A., Cleveland, W. S., & Shyu, M. (1996). The visual design and control of trellis display. *Journal of Computational and Graphical Statistics*, 5, 123–155. <http://dx.doi.org/10.1080/10618600.1996.10474701>
- Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, 65, 676–696. <http://dx.doi.org/10.1177/0013164404272507>
- Browne, M. W. (1968). A comparison of factor analytic techniques. *Psychometrika*, 33, 267–334. <http://dx.doi.org/10.1007/BF02289327>
- Browne, M. W. (1972). Orthogonal rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, 25, 115–120. <http://dx.doi.org/10.1111/j.2044-8317.1972.tb00482.x>
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111–150. http://dx.doi.org/10.1207/S15327906MBR3601_05
- Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, 80, 796–846. <http://dx.doi.org/10.1111/j.1467-6494.2011.00749.x>
- Camp, B. H. (1932). The converse of Spearman's two-factor theorem. *Biometrika*, 24, 418–427. <http://dx.doi.org/10.1093/biomet/24.3.4.418>
- Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2016). Factor structure of the Wechsler Intelligence Scale for Children—fifth edition: Exploratory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*, 28, 975–986. <http://dx.doi.org/10.1037/pas0000238>
- Carroll, J. B. (1953). An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 18, 23–38. <http://dx.doi.org/10.1007/BF02289025>
- Carroll, J. B. (1957). Biquartimin criterion for rotation to oblique simple structure in factor analysis. *Science*, 126, 1114–1115. <http://dx.doi.org/10.1126/science.126.3283.1114>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571312>
- Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31, 33–42. <http://dx.doi.org/10.1007/BF02289455>
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, 33, 107–112. <http://dx.doi.org/10.1177/001316447303300111>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.
- Cox, G. M. (1939). The multiple factor theory in terms of common elements. *Psychometrika*, 4, 59–68. <http://dx.doi.org/10.1007/BF02287993>
- Cucina, J. M., & Howardson, G. N. (2017). Woodcock-Johnson-III, Kaufman Adolescent and Adult Intelligence Test (KAIT), Kaufman Assessment Battery for Children (KABC), and Differential Ability Scales (DAS) support Carroll but not Cattell-Horn. *Psychological Assessment*, 29, 1001–1015. <http://dx.doi.org/10.1037/pas0000389>
- Dombrowski, S. C. (2014). Investigating the structure of the WJ-III cognitive in early school age through two exploratory bifactor analysis procedures. *Journal of Psychoeducational Assessment*, 32, 434–494. <http://dx.doi.org/10.1177/0734282914530838>
- Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2017). Exploratory and hierarchical factor analysis of the WJ-IV Cognitive at school age. *Psychological Assessment*, 29, 394–407. <http://dx.doi.org/10.1037/pas0000350>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299. <http://dx.doi.org/10.1037/1082-989X.4.3.272>
- Garnett, J. M. (1919). General ability, cleverness and purpose. *British Journal of Psychology*, 9, 345–366. <http://dx.doi.org/10.1111/j.2044-8295.1919.tb00232.x>
- Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: G as superordinate or breadth factor? *Psychological Science Quarterly*, 50, 21–43.
- Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, 55, 57–68. <http://dx.doi.org/10.1016/j.intell.2016.01.006>
- Harman, H. H. (1954). Karl John Holzinger. *Psychometrika*, 19, 95–96. <http://dx.doi.org/10.1007/BF02289158>
- Hattori, M., Zhang, G., & Preacher, K. J. (2017). Multiple local solutions and geomin rotation. *Multivariate Behavioral Research*, 52, 720–731. <http://dx.doi.org/10.1080/00273171.2017.1361312>
- Holzinger, K. J. (1936). Recent research on unitary mental traits. *Journal of Personality*, 4, 335–343. <http://dx.doi.org/10.1111/j.1467-6494.1936.tb02038.x>
- Holzinger, K. J. (1945). Spearman as I knew him. *Psychometrika*, 10, 231–235. <http://dx.doi.org/10.1007/BF02288890>
- Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54. <http://dx.doi.org/10.1007/BF02287965>
- Horst, P. (1941). A non-graphical method for transforming an arbitrary factor matrix into a simple structure factor matrix. *Psychometrika*, 6, 79–99. <http://dx.doi.org/10.1007/BF02292176>
- Jennrich, R. I. (2002). A simple general method or oblique rotation. *Psychometrika*, 67, 7–19. <http://dx.doi.org/10.1007/BF02294706>
- Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, 76, 1–13. <http://dx.doi.org/10.1007/s11336-011-9218-4>
- Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, 77, 442–454. <http://dx.doi.org/10.1007/s11336-012-9269-1>
- Jennrich, R. I., & Bentler, P. M. (2013). Erratum to: Exploratory bi-factor analysis. *Psychometrika*, 78, 556. <http://dx.doi.org/10.1007/s11336-013-9346-0>
- Kristof, W. (1970). A theorem on the trace of certain matrix products and some applications. *Journal of Mathematical Psychology*, 7, 515–530. [http://dx.doi.org/10.1016/0022-2496\(70\)90037-4](http://dx.doi.org/10.1016/0022-2496(70)90037-4)
- Kshirsagar, A. (1959). Bartlett decomposition and Wishart distribution. *The Annals of Mathematical Statistics*, 30, 239–241. <http://dx.doi.org/10.1214/aoms/1177706379>
- Litalien, D., Morin, A. J., Gagné, M., Vallerand, R. J., Losier, G. F., & Ryan, R. M. (2017). Evidence of a continuum structure of academic self-determination: A two-study test using a bifactor-ESEM representation of academic motivation. *Contemporary Educational Psychology*, 51, 67–82. <http://dx.doi.org/10.1016/j.cedpsych.2017.06.010>
- Lorenzo-Seva, U., & ten Berge, J. M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, 2, 57–64. <http://dx.doi.org/10.1027/1614-2241.2.2.57>
- MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, 36, 611–637. http://dx.doi.org/10.1207/S15327906MBR3604_06
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84–99. <http://dx.doi.org/10.1037/1082-989X.4.1.84>
- Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, 51, 698–717. <http://dx.doi.org/10.1080/00273171.2016.1215898>
- Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence*, 61, 120–129. <http://dx.doi.org/10.1016/j.intell.2017.01.012>

- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach. *Journal of Personality and Social Psychology*, 88, 139–157. <http://dx.doi.org/10.1037/0022-3514.88.1.139>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- 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, 149–161. <http://dx.doi.org/10.1080/00273171.2014.973990>
- Mulaik, S. A., & Quartetti, D. A. (1997). First order or higher order general factor? *Structural Equation Modeling*, 4, 193–211. <http://dx.doi.org/10.1080/10705519709540071>
- Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, 41, 407–422. <http://dx.doi.org/10.1016/j.intell.2013.06.004>
- Piaggio, H. T. H. (1931). The general factor in Spearman's theory of intelligence. *Nature*, 127, 56–67. <http://dx.doi.org/10.1038/127056b0>
- Pinneau, S. R., & Newhouse, A. (1964). Measures of invariance and comparability in factor analysis for fixed variables. *Psychometrika*, 29, 271–281. <http://dx.doi.org/10.1007/BF02289724>
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667–696. <http://dx.doi.org/10.1080/00273171.2012.715555>
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559. <http://dx.doi.org/10.1080/00223891.2010.496477>
- Reise, S., 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, 684–711. <http://dx.doi.org/10.1177/0013164410378690>
- Revelle, W., & Wilt, J. (2013). The general factor of personality: A critique. *Journal of Research in Personality*, 47, 493–504. <http://dx.doi.org/10.1016/j.jrp.2013.04.012>
- Rozeboom, W. W. (1992). The glory of suboptimal factor rotation: Why local minima in analytic optimization of simple structure are more blessing than curse. *Multivariate Behavioral Research*, 27, 585–599. http://dx.doi.org/10.1207/s15327906mbr2704_5
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53–61. <http://dx.doi.org/10.1007/BF02289209>
- Schönemann, P. H. (1966). A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31, 1–10. <http://dx.doi.org/10.1007/BF02289451>
- Serra, F., Spoto, A., Ghisi, M., & Vidotto, G. (2017). Improving major depressive episode assessment: A new tool developed by formal psychological assessment. *Frontiers in Psychology*. Advance online publication. <http://dx.doi.org/10.3389/fpsyg.2017.00214>
- Sharp, C., Wright, A. G. C., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors? *Journal of Abnormal Psychology*, 124, 387–398. <http://dx.doi.org/10.1037/abn0000033>
- Shen, W., Kiger, T. B., Davies, S. E., Rasch, R. L., Simon, K. M., & Ones, D. S. (2011). Samples in applied psychology: Over a decade of research in review. *Journal of Applied Psychology*, 96, 1055–1064. <http://dx.doi.org/10.1037/a0023322>
- Spearman, C. (1904). General intelligence' objectively determined and measured. *American Journal of Psychology*, 15, 201–293. <http://dx.doi.org/10.2307/1412107>
- Spearman, C. (1927). *The abilities of man*. New York, NY: Macmillan.
- Spearman, C., & Holzinger, K. J. (1934). *Preliminary reports on Spearman-Holzing Unitary Trait Study*. Chicago, IL: Statistical Laboratory, Department of Education, University of Chicago.
- Steiger, J. H. (1979). Factor indeterminacy in the 1930s and 1970s: Some interesting parallels. *Psychometrika*, 44, 157–167. <http://dx.doi.org/10.1007/BF02293967>
- Swanson, H. L. (2017). Verbal and visual-spatial working memory: What develops over a life span? *Developmental Psychology*, 53, 971–995. <http://dx.doi.org/10.1037/dev0000291>
- Thomson, G. H. (1951). *The factorial analysis of human ability*. New York, NY: Houghton Mifflin. (Original work published 1946)
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1–32. <http://dx.doi.org/10.1037/h0075959>
- Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits*. Chicago, IL: University of Chicago Press. <http://dx.doi.org/10.1037/10018-000>
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.
- Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago, IL: University of Chicago Press.
- Tucker, L. R. (1940). The role of correlated factors in factor analysis. *Psychometrika*, 5, 141–152. <http://dx.doi.org/10.1007/BF02287872>
- Tucker, L. R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army. <http://dx.doi.org/10.21236/AD0047524>
- Wainer, H. H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, 83, 312–317. <http://dx.doi.org/10.1037/0033-2909.83.2.213>
- Wainer, H. H. (1978). On the sensitivity of regression and regressors. *Psychological Bulletin*, 85, 267–273. <http://dx.doi.org/10.1037/0033-2909.85.2.267>
- Waller, N. G. (2018a). Direct Schmid–Leiman transformations and rank-deficient loadings matrices. *Psychometrika*, 83, 858–870. <http://dx.doi.org/10.1007/s11336-017-9599-0>
- Waller, N. G. (2018b). *fungible: Psychometric functions from the Waller Lab*. R package version 1.75.
- Waller, N. G. (2018c). An introduction to Kristof's theorem for solving least-square optimization problems without calculus. *Multivariate Behavioral Research*, 53, 190–198. <http://dx.doi.org/10.1080/00273171.2017.1412294>
- Wherry, R. J. (1959). Hierarchical factor solutions without rotation. *Psychometrika*, 24, 45–51. <http://dx.doi.org/10.1007/BF02289762>
- Wilks, S. S. (1938). Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika*, 3, 23–40. <http://dx.doi.org/10.1007/BF02287917>
- Wilson, E. B. (1928). Review of "The abilities of man, their nature and measurement" by C. Spearman. *Science*, 67, 244–248. <http://dx.doi.org/10.1126/science.67.1731.244>
- Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany, NY: SUNY Press.
- Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, 64, 113–128. <http://dx.doi.org/10.1007/BF02294531>

Received October 12, 2018

Revision received April 2, 2019

Accepted April 8, 2019 ■