Multivariate Behavioral Research, 45:104–134, 2010

ISSN: 0027-3171 print/1532-7906 online
DOI: 10.1080/00273170903504836

Psychology Press
Taylor & Francis Group

Bootstrap Confidence Intervals for Ordinary Least Squares Factor Loadings and Correlations in Exploratory Factor Analysis

Guangjian Zhang
University of Notre Dame

Kristopher J. Preacher
University of Kansas

Shanhong Luo
University of North Carolina at Wilmington

This article is concerned with using the bootstrap to assign confidence intervals for rotated factor loadings and factor correlations in ordinary least squares exploratory factor analysis. Coverage performances of *SE*-based intervals, percentile intervals, bias-corrected percentile intervals, bias-corrected accelerated percentile intervals, and hybrid intervals are explored using simulation studies involving different sample sizes, perfect and imperfect models, and normal and elliptical data. The bootstrap confidence intervals are also illustrated using a personality data set of 537 Chinese men. The results suggest that the bootstrap is an effective method for assigning confidence intervals at moderately large sample sizes.

Exploratory factor analysis (EFA) is widely used in the social sciences (Cudeck & MacCallum, 2007; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Gorsuch, 1983; Woods & Edwards, 2008; Yanai & Ichikawa, 2007). It allows the researcher to study unobservable constructs (e.g., personality traits) through observable measures (e.g., questionnaire items). The unobservable constructs are

Correspondence concerning this article should be addressed to Guangjian Zhang, Psychology Department, Haggar Hall, University of Notre Dame, Notre Dame, IN 46556. E-mail: gzhang3@nd.edu

referred to as latent factors; the directly measured variables are referred to as manifest variables.

Factor loadings of manifest variables on latent factors reveal the nature of latent factors. Factor loadings are not unique in EFA, however. Factor rotation (Browne, 2001) is conducted in EFA to achieve interpretability and identification. Two kinds of rotation are available: orthogonal rotation and oblique rotation. The factors are uncorrelated in orthogonal rotation and they can be correlated in oblique rotation. Oblique factors are arguably more realistic than orthogonal factors.

Interpretation of EFA results involves determining whether a particular manifest variable loads on a certain factor. A rule of thumb often used in practice is to ignore factor loadings whose absolute values are lower than 0.3. Cudeck and O'Dell (1994) criticized the blind use of this rule of thumb because of its inflexibility. They suggested that researchers consider standard error (*SE*) estimates or confidence intervals (CI).

When the EFA model is estimated using maximum likelihood (ML), CIs for rotated factor loadings and factor correlations can be constructed from point estimates and their *SE* estimates (see Jennrich, 1973b, for normal data and Yuan, Marshall, & Bentler, 2002, for nonnormal data, missing values, and outliers). Though maximum likelihood is a popular method of estimating the parameters of EFA, researchers are also interested in ordinary least squares (OLS) estimation. MacCallum and colleagues (Briggs & MacCallum, 2003; MacCallum, Browne, & Cai, 2007; MacCallum, Tucker, & Briggs, 2001) report that OLS estimation recovers weak factors more faithfully than maximum likelihood estimation when the factor analysis model contains model error in the population. Unlike maximum likelihood estimation, *SE* and CIs for OLS estimation have been unavailable until recently. Browne, Cudeck, Tateneni, and Mels (2008) implemented sandwich-type *SE* estimates (Browne, 1984, Proposition 2) of rotated factor loadings and factor correlations in OLS estimation. The *SE* estimates can be used to construct CIs when the manifest variables are normally distributed and the sample size is large. The *SE* estimates are asymptotically valid when the model holds exactly and when lack of fit is moderate (Browne, 1984, R5).

The goal of this article is to use the bootstrap to construct CIs for rotated factor loadings and factor correlations when the estimation method is OLS. The bootstrap (Efron & Tibshirani, 1993) is a nonparametric method for estimating *SEs*, constructing CIs, and conducting hypothesis tests. The bootstrap requires fewer assumptions than the conventional large-sample methods. The bootstrap has been used to estimate *SEs* for unrotated factor loadings (Chatterjee, 1984) and to construct approximate CIs for rotated factor loadings (Lambert, Wildt, & Durand, 1991). To the best of our knowledge, there were only two studies reporting simulation results for the bootstrap in EFA: one focuses on orthogonally rotated ML factor loading estimates (Ichikawa & Konishi, 1995) and the other one focuses on ML communality estimates (Ichikawa & Konishi, 2008). In both simulations, population manifest variable correlation matrices exactly satisfy a factor analysis model.

This article differs from previous works in three ways. First, this study concerns how bootstrap CIs perform for OLS estimation. Though the principal (common) factor method considered by Chatterjee (1984) and by Lambert et al. (1991) is similar to OLS estimation, the former is a noniterative procedure with squared multiple correlations as communality estimates. In practice, OLS estimation of EFA parameters is often implemented using an iterative procedure. Also, no simulation studies investigating OLS *SEs* or CIs have been reported to the best of our knowledge. Second, this study examines the performance of bootstrap CIs in oblique rotation: all previous work on the bootstrap in EFA considered only orthogonal rotation. Oblique rotation is preferred in most situations because it tends to give a clearer factor pattern and allows correlated factors. Third, this study examines how bootstrap CIs perform when the factor analysis model is misspecified in the population. It has been argued that all statistical models are wrong to some extent (Box, 1976; MacCallum, 2003). Statistical properties of a procedure under an imperfect model are directly relevant to the real world.

The remainder of the article is organized as follows: The next section briefly describes the EFA model and its estimation. A difficulty in using the bootstrap in EFA is that the rotated factor loading matrix is identified up to column reflection and column interchange. We adopt a method for reflecting and interchanging columns of rotated factor loading matrices to match the factor loading matrix of the parent sample. The effectiveness of the method is tested using a simulation study, followed by a section presenting how to construct bootstrap CIs using bootstrap replications of factor loadings and factor correlations. Simulation studies are employed to explore the influence of model error, manifest variable distributions, and sample size on the bootstrap CIs. In addition, the bootstrap CIs are illustrated using an empirical data set. Concluding comments are provided at the end of the article.

EXPLORATORY FACTOR ANALYSIS

The factor analysis model specifies that p manifest variables $\boldsymbol{y}$ are linear combinations of m latent factors $\boldsymbol{z}$ and p unique variables $\boldsymbol{u}$,

$$\boldsymbol{y} = \boldsymbol{\Lambda} \boldsymbol{z} + \boldsymbol{u}. \tag{1}$$

Here, $\boldsymbol{\Lambda}$ is a $p \times m$ factor loading matrix representing the influence of $\boldsymbol{z}$ on $\boldsymbol{y}$. Equation 1 is often referred to as the factor analysis data model. Under the assumptions that the common factors and the unique factors are uncorrelated and the unique factors are mutually uncorrelated, Equation 1 implies a correlation structure,

$$\boldsymbol{P} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{D}_{\psi}. \tag{2}$$

Here, $\boldsymbol{P}$ is the manifest variable correlation matrix, $\boldsymbol{\Phi}$ is the factor correlation matrix, and $\boldsymbol{D}_{\psi}$ is a diagonal matrix of unique variances.

When the EFA model is estimated in practice, the population manifest variable correlation matrix $\boldsymbol{P}$ is replaced by the sample manifest variable correlation matrix $\boldsymbol{R}$. It is often convenient to first assume that the factors are uncorrelated. Thus Equation 2 is reduced to[1]

$$\boldsymbol{P} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{D}_{\psi}. \tag{3}$$

Estimates of the factor loading matrix $\boldsymbol{\Lambda}$ and the matrix of unique variances $\boldsymbol{D}_{\psi}$ can be obtained by minimizing the maximum likelihood discrepancy function,[2]

$$f_{\mathrm{ml}} = \ln|\hat{\boldsymbol{P}}| + \mathrm{trace}(\boldsymbol{R}\hat{\boldsymbol{P}}^{-1}) - \ln|\boldsymbol{R}| - p, \tag{4}$$

or the OLS discrepancy function,

$$f_{\mathrm{ols}} = \mathrm{trace}(\boldsymbol{R} - \hat{\boldsymbol{P}})^2. \tag{5}$$

Minimization of Equations 4 and 5 with regard to $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$ requires an iterative algorithm. The iterative algorithm may not converge in the region where all unique variances are positive. One way to improve convergence is to impose the constraints that unique variance estimates must be positive. When a unique variance estimate is on the boundary of 0, a Heywood case is said to have occurred (Lawley & Maxwell, 1971, p. 32).

The EFA model of Equation 3 is not identified. In particular, it still holds if the factor loading matrix $\boldsymbol{\Lambda}$ is postmultiplied by an orthogonal matrix. This is referred to as rotational indeterminacy. Rotational indeterminacy can be removed by imposing $m(m-1)/2$ constraints to the factor loading matrix $\boldsymbol{\Lambda}$. For example, ML estimation requires that

$$\boldsymbol{\Lambda}'\boldsymbol{D}_{\psi}^{-1}\boldsymbol{\Lambda} = \boldsymbol{D}_1, \tag{6}$$

[1]A reviewer points out that the factor analysis model can be considered as partitioning the manifest variable correlation matrix $\boldsymbol{P}$ into two parts: one part due to the influence of factors $\boldsymbol{P}^t = \boldsymbol{P} - \boldsymbol{D}_{\psi}$ and the other part due to the influence of unique variables $\boldsymbol{D}_{\psi}$. In EFA, both $\boldsymbol{P}^t$ and $\boldsymbol{D}_{\psi}$ are uniquely defined but $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$ are not.

[2]To be precise, the sample correlation matrix, $\boldsymbol{R}$, and model implied correlation matrix, $\hat{\boldsymbol{P}}$, should be replaced by the sample covariance matrix, $\boldsymbol{S}$, and model implied covariance matrix, $\hat{\boldsymbol{\Sigma}}$ (Cudeck, 1989). However, EFA models are scale-invariant. Thus, fitting the model to the sample correlation matrix yields correct results.

and OLS estimation requires that

$$\boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \boldsymbol{D}_2. \tag{7}$$

Here, $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are diagonal matrices.[3]

The two identification conditions in Equations 6 and 7 are motivated by mathematical convenience rather than meaningful interpretation. Interpretation of factor analysis is dramatically improved by factor rotation (Browne, 2001),

$$\boldsymbol{\Lambda}_{\mathrm{r}} = \boldsymbol{\Lambda}\boldsymbol{T}. \tag{8}$$

Here, $\boldsymbol{\Lambda}_{\boldsymbol{r}}$ is the rotated factor matrix and $\boldsymbol{T}$ is a transformation matrix. The transformation matrix $\boldsymbol{T}$ also produces the factor correlation matrix,

$$\boldsymbol{\Phi} = (\boldsymbol{T}'\boldsymbol{T})^{-1}. \tag{9}$$

In orthogonal rotation, the transformation matrix $\boldsymbol{T}$ is an orthogonal matrix and the resulting factor correlation matrix $\boldsymbol{\Phi}$ is the identity matrix. In oblique rotation, $\boldsymbol{\Phi}$ has unit diagonal elements and nonzero off-diagonal elements. The transformation matrix $\boldsymbol{T}$ is determined using certain rotation criteria (Browne, 2001).

A versatile rotation criterion is the Crawford-Ferguson family, which is a sum of two terms (Crawford & Ferguson, 1970, Equation 7),

$$f(\boldsymbol{\Lambda}) = (1-\kappa)\sum_{i=1}^{p}\sum_{j=1}^{m}\sum_{l\neq j}^{m}\lambda_{ij}^2\lambda_{il}^2 + \kappa\sum_{j=1}^{m}\sum_{i=1}^{p}\sum_{k\neq i}^{p}\lambda_{ij}^2\lambda_{kj}^2. \tag{10}$$

The first term measures row complexity and attains its minimum value of zero if every row has at most one nonzero element. The second term measures column complexity and attains its minimum value of zero if every column has at most one nonzero element. Relative contributions of row complexity and column complexity are determined by κ. The parameter κ can take any value between 0 and 1. Many popular rotation criteria can be regarded as special cases of the Crawford-Ferguson family. For example, when orthogonal rotation is considered, minimizing the Crawford-Ferguson family with $\kappa = 1/p$ and maximizing the varimax criterion (Kaiser, 1958) give exactly the same rotated factor loading matrix. Thus the Crawford-Ferguson family with $\kappa = 1/p$ is referred to as *CF-varimax* rotation. Though varimax rotation tends to fail for oblique rotation, CF-varimax rotation often gives satisfactory results for oblique rotation (Browne, 2001, Table 8).

[3]The diagonal elements of $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ need to be ordered and distinct.

CIs for rotated factor loadings and factor correlations provide a range of plausible values for these parameters. This article focuses on how to construct bootstrap CIs for rotated factor loadings and factor correlations under OLS estimation. The bootstrap of EFA encounters an unusual difficulty: the rotated factor loading matrix $\mathbf{\Lambda}_r$ is identified up to column interchange and reflection (Jennrich, 2007; Pennel, 1972).

THE "ALIGNMENT PROBLEM" AND THE COLUMN REFLECTION AND INTERCHANGE METHOD

Table 1 presents three sets of factor loading matrices and factor correlation matrices obtained from fitting a three-factor model to cognitive ability data presented by Porac, Coren, Steiger, and Duncan (1980). Multiplying the first

TABLE 1
Factor Loading and Factor Correlation Matrices to Illustrate Column Interchange and Reflection

	$\mathbf{\Lambda}_1$			$\mathbf{\Lambda}_2$			$\mathbf{\Lambda}_3$		
	F1	*F2*	*F3*	*F1*	*F2*	*F3*	*F1*	*F2*	*F3*
Ball	0.851	0.054	0.049	−0.851	0.054	0.049	0.049	0.054	0.851
Draw	0.861	0.061	−0.001	−0.861	0.061	−0.001	−0.001	0.061	0.861
Eraser	0.854	0.037	0.020	−0.854	0.037	0.020	0.020	0.037	0.854
Card	0.442	0.059	0.091	−0.442	0.059	0.091	0.091	0.059	0.442
Kick	0.623	0.103	0.137	−0.623	0.103	0.137	0.137	0.103	0.623
Pebble	0.434	0.072	0.186	−0.434	0.072	0.186	0.186	0.072	0.434
Chair	0.309	0.097	0.226	−0.309	0.097	0.226	0.226	0.097	0.309
Keyhole	−0.036	0.914	0.017	0.036	0.914	0.017	0.017	0.914	−0.036
Bottle	−0.005	0.910	0.015	0.005	0.910	0.015	0.015	0.910	−0.005
Rifle	0.114	0.623	0.039	−0.114	0.623	0.039	0.039	0.623	0.114
Door	−0.064	0.015	0.891	0.064	0.015	0.891	0.891	0.015	−0.064
Heart	−0.025	0.020	0.801	0.025	0.020	0.801	0.801	0.020	−0.025
Earphone	0.131	0.009	0.550	−0.131	0.009	0.550	0.550	0.009	0.131
	$\mathbf{\Phi}_1$			$\mathbf{\Phi}_2$			$\mathbf{\Phi}_3$		
	F1	*F2*	*F3*	*F1*	*F2*	*F3*	*F1*	*F2*	*F3*
F1	1.000			1.000			1.000		
F2	0.278	1.000		−0.278	1.000		0.268	1.000	
F3	0.234	0.268	1.000	−0.234	0.268	1.000	0.234	0.278	1.000

Note. Factor loadings and correlations were obtained from ordinary least squares (OLS) factor analysis of Porac, Coren, Steiger, & Duncan (1980) cognitive ability data.

column of $\mathbf{\Lambda}_1$ by -1 produces $\mathbf{\Lambda}_2$; Interchanging the first and third columns of $\mathbf{\Lambda}_1$ produces $\mathbf{\Lambda}_3$. Because most rotation criteria[4] are defined on squared factor loadings, the three factor loading matrices will yield the same rotation criterion function value.

The column reflection and interchange of the factor loading matrix can be expressed algebraically. Let

$$\boldsymbol{W} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

be a reflection matrix: it is obtained by changing the sign of the first column of an identity matrix. The factor loading matrix $\mathbf{\Lambda}_2$ is given by

$$\mathbf{\Lambda}_2 = \mathbf{\Lambda}_1 \boldsymbol{W}, \tag{11}$$

and the corresponding factor correlation matrix is given by

$$\mathbf{\Phi}_2 = \boldsymbol{W}' \mathbf{\Phi}_1 \boldsymbol{W}. \tag{12}$$

Similarly, let

$$\boldsymbol{W} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

be an interchange matrix: it is obtained by interchanging the first and third columns of an identity matrix. Manipulations similar to Equations 11 and 12 yield the factor loading matrix $\mathbf{\Lambda}_3$ and factor correlation matrix $\mathbf{\Phi}_3$.

Column interchange and reflection produce a large number of factor matrices with the same rotation criterion function value. Let k be the number of factors. The total number of such factor matrices is $2^k k!$ (Ichikawa & Konishi, 1995): it is 8 for a two-factor model, 48 for a three-factor model, 384 for a four-factor model, and 3,840 for a five-factor model.

All these $2^k k!$ factor loading matrices are in fact one solution in disguise. Failing to recognize this leads to inflated bootstrap *SE* estimates and erroneous bootstrap CIs. This issue has been referred to as the "alignment problem" by several authors (Clarkson, 1979; Pennel, 1972).

To deal with the "alignment problem," we reflect and interchange columns of rotated factor loading matrices $\mathbf{\Lambda}^*$ in bootstrap samples to match an order matrix $\mathbf{\Lambda}$. The order matrix $\mathbf{\Lambda}$ is the rotated factor loading matrix of the parent

[4] An exception is target rotation, in which the factor matrix is rotated to match a partially specified target matrix (see Browne, 2001, pp. 123–125.

sample. The match is conducted as follows (Section 2.1.3., Brown et al., 2008; Clarkson, 1979; Ichikawa & Konishi, 1995). Column reflection and interchange produces $2^k k!$ rotated factor loading matrices $\mathbf{\Lambda}^*$ in each bootstrap sample. We define the sum of squared differences between elements of an arbitrary $\mathbf{\Lambda}^*$ and the corresponding elements of the order matrix $\mathbf{\Lambda}$ as

$$f(\mathbf{\Lambda}; \mathbf{\Lambda}^*) = \sum_{j=1}^{m} \sum_{i=1}^{p} (\lambda_{ij}^* - \lambda_{ij})^2. \tag{13}$$

The $\mathbf{\Lambda}^*$ that produces the minimum value of $f(\mathbf{\Lambda}; \mathbf{\Lambda}^*)$ among the $2^k k!$ $\mathbf{\Lambda}^*$s in a bootstrap sample is said to match the order matrix $\mathbf{\Lambda}$. Though the sign or ordering of elements of the rotated bootstrap factor loading matrix $\mathbf{\Lambda}^*$ may change, their absolute values remain the same. After columns of the bootstrap factor loading matrix $\mathbf{\Lambda}^*$ are properly reflected and interchanged, the bootstrap factor correlation matrix is changed according to Equation 12. R code reflecting and interchanging columns of bootstrap factor loading matrices to match an order matrix is available on the Internet (http://www.nd.edu/∼gzhang3/Papers/BootEFA/AlignmentRcode.txt).

The alignment problem becomes more demanding as the number of factors becomes larger. For example, Jennrich (2007, p. 332) expressed concerns of the "devastating" effects of alignment failures on *SE* estimates of the bootstrap. In the next section, we examine the alignment problem more closely and suggest a diagnostic tool for alignment failures.

THE EFFECTIVENESS OF THE COLUMN REFLECTION AND INTERCHANGE METHOD

To examine the effectiveness of the column reflection and interchange algorithm, we conduct a simulation study in which factor loading matrices of simulation samples are reflected and interchanged to match the population factor loading matrix. Standard deviations of factor loadings of simulation samples are then compared with asymptotic *SEs*.

Population and Procedures

The population factor loading matrix we used was originally constructed by Jennrich (1973a, Table 2) to demonstrate the influence of factor rotation on *SEs* of factor loadings. Table 2 shows the unrotated factor loading matrix and two orthogonally rotated factor loading matrices. Jennrich (1973a) deliberately constructed the artificial data set so the unrotated factor loading matrix and the

TABLE 2
Factor Loading Matrices of an Artificial Example Presented by Jennrich (1973a)

Unrotated		*Varimax*		*Crawford-Ferguson*	
0.81 (0.06)	0.00 (0.09)	0.81 (0.06)	0.00 (2.34)	−0.57 (0.13)	−0.57 (0.13)
0.81 (0.06)	0.00 (0.09)	0.81 (0.06)	0.00 (2.34)	−0.57 (0.13)	−0.57 (0.13)
−0.43 (0.11)	0.43 (0.12)	−0.43 (1.28)	0.43 (1.32)	0.61 (0.10)	0.00 (0.13)
−0.43 (0.11)	0.43 (0.12)	−0.43 (1.28)	0.43 (1.32)	0.61 (0.10)	0.00 (0.13)
−0.43 (0.11)	−0.43 (0.12)	−0.43 (1.28)	−0.43 (1.32)	0.00 (0.13)	0.61 (0.10)
−0.43 (0.11)	−0.43 (0.12)	−0.43 (1.28)	−0.43 (1.32)	0.00 (0.13)	0.61 (0.10)

Note. The asymptotic *SEs* scaled for $n = 100$ are presented in parentheses.

varimax rotated factor loading matrix are identical. Though the unrotated factor loading matrix is well identified, the varimax rotated factor loading matrix is nearly unidentified. This population factor loading matrix was suggested by a reviewer because it violates Thurstone's rules for simple structure. In addition, we consider the Crawford-Ferguson rotation with $\kappa = 0.5$. To make the results comparable to those reported in Jennrich (1973a, Table 2), we also extract factors using ML and consider orthogonal rotation in this simulation study.

Also shown in Table 2 are asymptotic *SEs* of the unrotated and the rotated factor loadings. These asymptotic *SEs* are scaled at the sample size $n = 100$. The asymptotic *SEs* of rotated factor loadings were computed using CEFA 3.02 (Browne et al., 2008). The asymptotic *SEs* of the unrotated factor loadings and varimax rotated factor loadings were also provided in Jennrich (1973a, Table 2).

Simulation samples were generated according to the factor analysis model of Equation 1: the factor loadings are presented in Table 2; the two factors follow two independent standard normal distributions; the six unique variables follow independent normal distribution with variances of 0.34, 0.34, 0.63, 0.63, 0.63, and 0.63. Thus the six manifest variables are also normally distributed. Two sample sizes are considered: the sample size $n = 100$ is used to explore the performance of the asymptotic *SEs* and empirical *SEs* in small samples; the sample size $n = 1{,}000$ is used to explore their performance in large samples. At each sample size, $k = 1{,}000$ samples are simulated. In each simulated sample, two factors are extracted using ML. Two separate orthogonal rotations are conducted: the CF-varimax rotation and Crawford-Ferguson with $\kappa = 0.5$. The unrotated factor loading matrix, the varimax rotated factor loading matrix, and the Crawford-Ferguson rotated factor loading matrix in each simulated sample are aligned to the corresponding population factor loading matrices presented in Table 2. Standard deviations of factor loadings of $k = 1{,}000$ simulation samples are referred to as empirical *SEs*.

Results

The empirical *SEs* and the asymptotic *SEs* were expected to be close to each other at the sample size of $n = 1{,}000$ because the sample size is large and the normality assumption is true. We computed ratios between empirical *SEs* and asymptotic *SEs* to examine their similarity. A ratio of 1.00 denotes perfect similarity and thus perfect column reflection and interchange. Figure 1 displays the empirical *SE*/asymptotic *SE* ratios for the unrotated factor loadings, varimax rotated factor loadings, and Crawford-Ferguson rotated factor loadings. At the sample size of $n = 1{,}000$, asymptotic *SEs* and empirical *SEs* are close to each other for both the unrotated factor loadings and Crawford-Ferguson rotated factor loadings. Two conclusions seem reasonable in these two conditions. First, the column reflection and interchange algorithm works well. Second, the sample size $n = 1{,}000$ seems sufficiently large for the asymptotic *SE* estimates to give close approximations.

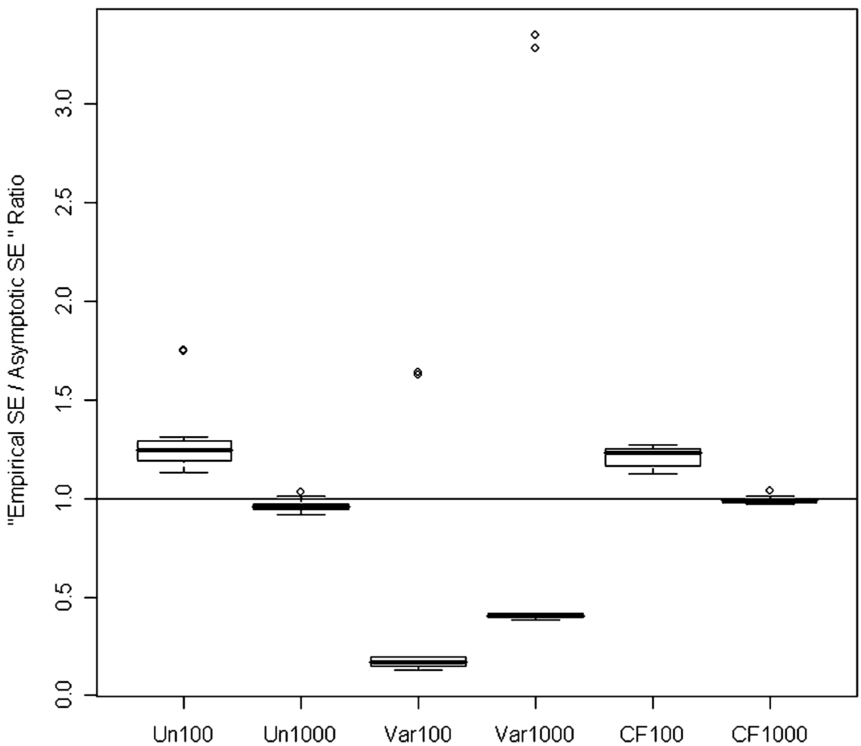

FIGURE 1 Ratios of empirical *SEs* to asymptotic *SEs* for three rotation methods."Un" = unrotated; "Var" = varimax; "CF" = Crawford-Ferguson $\kappa = 0.5$.

At the sample size of $n = 100$, the empirical *SEs* are slightly larger than the asymptotic *SEs* for unrotated factor loadings and Crawford-Ferguson rotated factor loadings. Two explanations of these discrepancies are possible. First, the column reflection and interchange method did not work satisfactorily. To examine the performance of the column reflection and interchange algorithm, we take a closer look at sample estimates of factor loadings from Table 2. The procedure is as follows: First, m rows of $\mathbf{\Lambda}$, each with a single large factor loading, are selected. Then histograms are produced for sample estimates of each of these large loadings. If columns are poorly aligned or reflected, then histograms of the selected loadings will appear bimodal or will have outliers. The left panel of Figure 2 displays the histograms of sample estimates of the unrotated factor loading λ_{11} at $n = 100$ and $n = 1{,}000$. The column reflection and interchange algorithm seems to work well even at $n = 100$ because sample estimates follow a unimodal distribution and the smallest value is 0.41.

The second possible reason for the discrepancies between the asymptotic and empirical *SEs* of unrotated and Crawford-Ferguson $\kappa = 0.5$ rotated factor

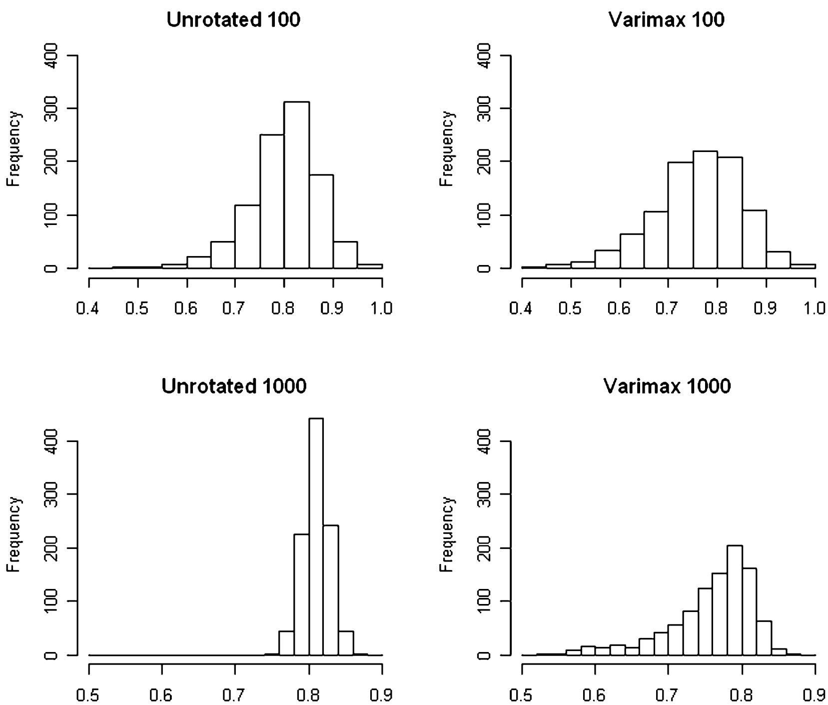

FIGURE 2 Distributions of factor loadings from simulated samples drawn from Jennrich's (1973a) example.

loadings at $n = 100$ is that this sample size may not be large enough for satisfactory asymptotic *SEs*. This conjecture is in accordance with a previous finding that asymptotic *SE* estimates were slightly smaller than exact ones at the sample size of $n = 150$ when manifest variables were normally distributed (Ichikawa & Konishi, 1995, Table 3, the columns corresponding to $\epsilon = 0.0$).

As shown in Figure 1, the empirical *SEs* and the asymptotic *SEs* differ substantially from each other for the varimax rotated factor loadings at both sample sizes of $n = 100$ and $n = 1{,}000$. In particular, empirical *SEs* are much smaller than asymptotic *SEs* for most factor loadings and empirical *SEs* are much larger than the corresponding asymptotic *SEs* for two-factor loadings. The right panel of Figure 2 displays the histograms of sample estimates of the varimax rotated factor loading λ_{11} at $n = 100$ and $n = 1{,}000$. The sample estimates follow a unimodal distribution in both conditions. The smallest value is 0.42 at $n = 100$ and 0.53 at $n = 1{,}000$. The histograms suggest that the column reflection and interchange method works well in both conditions. The discrepancies between asymptotic *SEs* and empirical *SEs* are likely due to the fact that the sample sizes are not large enough for asymptotic theory to give a close approximation.[5]

Discussion

The simulation study shows that the column reflection and interchange method works satisfactorily under ideal conditions of normally distributed manifest variables, a well identified model, and large sample sizes. The column reflection and interchange method seems to work well in less optimal conditions where the sample size is small. Though the column reflection and interchange algorithm appeared to work satisfactorily in this simulation study, generalization of the results to other conditions should be made with caution. It is desirable to detect alignment failures. The procedure described earlier is one effective strategy for detecting poor column reflection and/or interchange. With confidence in the column reflection and interchange method established, we next describe how to construct bootstrap CIs for rotated factor loadings and factor correlations.

BOOTSTRAP CIs

Bootstrap CIs are constructed using bootstrap replications of parameters and point estimates of the parent sample. We first describe how to obtain bootstrap

[5]The empirical *SE* estimates and the asymptotic *SE* estimates are much closer at the sample size of 1 million.

replications of rotated factor loadings and factor correlations. We then present methods for assigning bootstrap CIs.

Bootstrap Replications of Factor Loadings and Factor Correlations

The bootstrap replications are obtained using a three-step procedure. In the first step, the EFA model is fitted to the correlation matrix $\boldsymbol{R}$ of the parent sample, $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$. The estimation method is OLS. Let $\hat{\boldsymbol{\Lambda}}$ and $\hat{\boldsymbol{\Phi}}$ be the rotated factor loading matrix and factor correlation matrix of the parent sample.[6]

In the second step, a bootstrap sample $\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \ldots, \boldsymbol{y}_n^*$ is drawn from the parent sample $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$ with replacement.

In the third step, a correlation matrix $\boldsymbol{R}^*$ is computed from the bootstrap sample $\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \ldots, \boldsymbol{y}_n^*$. The EFA model is fitted to the bootstrap sample correlation matrix $\boldsymbol{R}^*$ using OLS estimation. The rotation method is the same as the one considered in the first step. Let $\hat{\boldsymbol{\Lambda}}^*$ and $\hat{\boldsymbol{\Phi}}^*$ be the corresponding rotated factor loading matrix and factor correlation matrix. It is critical that columns of the bootstrap factor loading matrix $\hat{\boldsymbol{\Lambda}}^*$ are reflected and interchanged to match the parent factor loading matrix $\hat{\boldsymbol{\Lambda}}$ as in Equation 13. Columns and rows of the bootstrap factor correlation matrix $\hat{\boldsymbol{\Phi}}^*$ are changed accordingly.

Repeating the second and third steps B times produces B bootstrap factor loading matrices $\hat{\boldsymbol{\Lambda}}_1^*, \hat{\boldsymbol{\Lambda}}_2^*, \ldots, \hat{\boldsymbol{\Lambda}}_B^*$, and B factor correlation matrices, $\hat{\boldsymbol{\Phi}}_1^*, \hat{\boldsymbol{\Phi}}_2^*, \ldots, \hat{\boldsymbol{\Phi}}_B^*$.

Let θ be a parameter of interest: it can be a rotated factor loading, a factor correlation; a unique variance; a communality; or even a function of parameters, for example, the difference between two factor loadings. We consider five different types of bootstrap CIs: the *SE*-based CI, the percentile interval, the bias-corrected percentile interval, the bias-corrected accelerated percentile interval, and a hybrid interval.

The *SE*-Based Bootstrap CI

The *SE*-based bootstrap CI is constructed using the point estimate $\hat{\theta}$ and the bootstrap *SE* estimate $\widehat{SE}_{\hat{\theta}}$. For example, the $(1-\alpha)$ *SE*-based CIs (Efron & Tibshirani, 1993, chap. 12),

$$(\hat{\theta}_{\text{lo}}, \hat{\theta}_{\text{up}}) = (\hat{\theta} - \widehat{SE}_{\hat{\theta}} \times x^{(1-\alpha/2)}, \hat{\theta} - \widehat{SE}_{\hat{\theta}} \times x^{(\alpha/2)}). \tag{14}$$

[6]Because the bootstrap involves only the rotated factor loading matrix $\boldsymbol{\Lambda}_{\text{r}}$, the subscript is dropped for ease of presentation.

Here $x^{(\alpha/2)}$ and $x^{(1-\alpha/2)}$ are critical points. A quick and easy choice of the critical points is the normal percentiles $z^{(\alpha/2)}$ and $z^{(1-\alpha/2)}$, for example, $z^{(.95)} = 1.645$. The bootstrap *SE* estimate $\widehat{SE}_{\hat{\theta}}$ is the standard deviation of the B bootstrap replications, $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$,

$$\widehat{SE}_{\hat{\theta}} = \left(\frac{\sum_{i=1}^{B} (\hat{\theta}_i^* - \overline{\theta}^*)^2}{B-1} \right)^{\frac{1}{2}}. \tag{15}$$

Here $\overline{\theta}^*$ is the mean of the B bootstrap replications.

SE-based bootstrap CIs will be constructed for factor loadings using Equations 14 and 15. If θ has natural bounds, for example, factor correlations are bounded by -1 and 1 and unique variances and communalities are bounded by 0 and 1, there is no guarantee that *SE*-based CIs for these parameters are within the natural bounds. One solution is to consider monotonic transformations for the parameters so transformed scales will be unbounded. CIs are constructed using the transformed scales first and then are transformed back to the original scales. The Fisher Z transformation is considered for factor correlations. A similar transformation is considered for unique variances and communalities (Browne, 1982, Equations 1.6.41, 1.6.42, and 1.6.43).

The Bootstrap Percentile Interval

The percentile interval (Efron & Tibshirani, 1993, p. 171) is constructed from percentiles of the B bootstrap replications $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$. Let $\hat{\theta}_b^{*(\beta)}$ be the βth bootstrap percentile. A percentile interval is given by

$$(\hat{\theta}_{\text{lo}},\ \hat{\theta}_{\text{up}}) = \left(\hat{\theta}_b^{*(\beta_1)},\ \ \hat{\theta}_b^{*(\beta_2)}\right) \tag{16}$$

where $0 < \beta_1 < \beta_2 < 1$. To construct a $(1-\alpha)$ bootstrap percentile interval, β_1 and β_2 are chosen to be $\alpha/2$ and $1-\alpha/2$, respectively.

The Bias-Corrected Accelerated Percentile Interval

A much more sophisticated method for choosing β_1 and β_2 in Equation 16 is the bias-corrected accelerated percentile interval (BCa; Efron & Tibshirani, 1993, pp. 184–188), which involves estimation of a bias parameter, z_0, and a parameter of acceleration, a,

$$\beta_1 = \Phi\left(z_0 + \frac{z_0 + z^{(\alpha/2)}}{1 - a(z_0 + z^{(\alpha/2)})}\right),$$

$$\beta_2 = \Phi\left(z_0 + \frac{z_0 + z^{(1-\alpha/2)}}{1 - a(z_0 + z^{(1-\alpha/2)})}\right). \tag{17}$$

Here $\boldsymbol{\Phi}(\cdot)$ is the cumulative distribution function of the standard normal distribution. The bias parameter z_0 is estimated from the proportion of bootstrap replications $\hat{\theta}^*$ less than the sample parameter estimate $\hat{\theta}$,

$$\hat{z}_0 = \boldsymbol{\Phi}^{-1}\left(\frac{\text{Numbers of bootstrap replications}\left\{\hat{\theta}^* < \hat{\theta}\right\}}{B}\right). \tag{18}$$

Here $\boldsymbol{\Phi}^{-1}(\cdot)$ is the inverse function of the standard normal distribution. The parameter of acceleration, a, is estimated using the jackknife method,

$$\hat{a} = \frac{\sum_{i=1}^{n}\left(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)}\right)^3}{6\left\{\sum_{i=1}^{n}\left(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)}\right)^2\right\}^{\frac{3}{2}}}. \tag{19}$$

Here $\hat{\theta}_{(i)}$ is the parameter estimate obtained from the ith jackknife sample, $\boldsymbol{y}_1$, $\boldsymbol{y}_2, \ldots, \boldsymbol{y}_{i-1}, \boldsymbol{y}_{i+1}, \ldots, \boldsymbol{y}_n$. The mean of $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \ldots, \hat{\theta}_{(n)}$ is $\hat{\theta}_{(\cdot)}$.

The Bias-Corrected Percentile Interval

The bias-corrected (BC) percentile interval can be regarded as a special case of the BCa of Equation 17, in which the acceleration parameter a is set to zero.

The Hybrid Interval

The hybrid interval (Hall, 1988; Ichikawa & Konishi, 1995) is constructed using the point estimate $\hat{\theta}$ and bootstrap percentiles,

$$(\hat{\theta}_{\text{lo}},\ \hat{\theta}_{\text{up}}) = \left(2\hat{\theta} - \hat{\theta}_b^{*(1-\alpha/2)},\ \ 2\hat{\theta} - \hat{\theta}_b^{*(\alpha/2)}\right). \tag{20}$$

It can be considered a modification of the *SE*-based CI of Equation 14. The critical points $x^{(1-\alpha/2)}$ and $x^{(\alpha/2)}$ are estimated using bootstrap percentiles,

$$x^{(1-\alpha/2)} = \frac{\hat{\theta}_b^{*(1-\alpha/2)}}{\hat{SE}_{\hat{\theta}}} \text{ and } x^{(\alpha/2)} = \frac{\hat{\theta}_b^{*(\alpha/2)}}{\hat{SE}_{\hat{\theta}}}. \tag{21}$$

EXPLORING PERFORMANCES OF BOOTSTRAP CIs

Simulation studies were conducted to explore the performance of bootstrap CIs of rotated factor loadings and factor correlations in a variety of conditions. The simulation studies involved two population conditions ("no model error" vs. "model error"), four sample sizes, and two manifest variable distributions (a normal distribution and an elliptical distribution).

Populations

A 13×13 manifest variable correlation matrix reported by Porac et al. (1980) is used as a population manifest variable correlation matrix. The root mean square error of approximation (RMSEA, Brown & Cudeck, 1993) for the three-factor model is 0.075 with a 90% CI of $(0.067, 0.084)$. The rotated factor loading matrix and factor correlation matrix are shown in the left panel of Table 1. The three factors are "limb" (Factor 1), "eye" (Factor 2), and "ear" (Factor 3). The estimation method is OLS and the rotation criterion is oblique CF-varimax (Browne, 2001).

One of the goals of the simulation studies was to explore the influence of model error on bootstrap CIs. In half of the conditions, population manifest variable correlation matrices satisfy the factor analysis model exactly: they are constructed using the factor loading matrices and factor correlation matrices of Table 1 according to Equation 2. These conditions are referred to as "no model error" conditions. In the other half of the conditions, factor analysis models provide approximations to population manifest variable correlation matrices, which are the sample correlation matrix reported by Porac, Coren, Steiger, and Duncan (1980). These conditions are referred to as "model error" conditions. Examining the residuals obtained from fitting the factor analysis model to the population manifest variable correlation matrix suggests that the model error is like background noise.

Four sample sizes were considered: $n = 100, 200, 300$, and 962. The original sample size was $n = 962$, which is fairly large for psychological research. The sample sizes of $n = 100, 200$, and 300 were included because many applications of factor analysis in psychology have similar sample sizes.

No modifications are required for bootstrap CIs when manifest variables are not normally distributed. To demonstrate this point, manifest variables are generated from either a multivariate normal distribution or an elliptical distribution. In normal distribution conditions, manifest variables are generated from the multivariate normal distribution $N_p(\mathbf{0}, \boldsymbol{P})$ where $\boldsymbol{P}$ may or may not satisfy the factor analysis model. Under the elliptical distribution conditions, manifest variables are generated from a mixture of normal distributions (Ichikawa & Konishi, 1995),

$$0.7N_p(0, \boldsymbol{P}) + 0.3N_p(0, 3\boldsymbol{P}). \tag{22}$$

Note that this distribution is a multivariate normal distribution $N_p(0, \boldsymbol{P})$ contaminated with another multivariate normal distribution $N_p(0, 3\boldsymbol{P})$. The contamination rate is 30%. The multivariate normal distribution is also a member of the elliptical distribution family but it has thinner tails than this elliptical distribution.

Procedures

Five hundred samples ($k = 500$) are simulated in each of the 16 conditions (2 population conditions $\times$ 4 sample sizes $\times$ 2 manifest variable distributions). Each simulation sample is like a parent sample. We constructed the *SE*-based bootstrap CIs, percentile intervals, BC, BCa, and hybrid intervals for rotated factor loadings and factor correlations in each simulation sample. The nominal coverage level is 90% for all bootstrap CIs. The number of bootstrap samples is $B = 2{,}000$. The estimation method is OLS and the rotation criterion is oblique CF-varimax rotation (Browne, 2001).

A Fortran program was prepared to generate manifest variables from the normal distribution and the elliptical distribution to draw bootstrap samples from a simulation sample and to construct bootstrap CIs from bootstrap replications. Model estimation, factor rotation, and in particular column interchange and reflection were carried out using the program CEFA 3.02 (Browne et al., 2008).

Results

For the sake of simple presentation, we highlight only the most important results of the simulation study. More detailed results are provided in a separate supporting file, which can be downloaded from the Internet (http://www.nd.edu/~gzhang3/Papers/BootEFA/SupportingMaterials.pdf). Before reporting the coverage performances of CIs constructed in the simulation study, we first discuss how to deal with Heywood cases.

TABLE 3
Frequencies of Heywood Cases in Simulation Samples

Sample size	Model Error		No Model Error	
	Normal	*Elliptical*	*Normal*	*Elliptical*
n = 100	36	47	23	40
n = 200	6	6	3	2
n = 300	1	3	1	2
n = 962	0	0	0	0

Note. Each condition involves 500 simulation samples.

Heywood cases. The factor analysis model converged in all simulation samples and bootstrap samples, but Heywood cases did occur in a number of simulation samples and bootstrap samples. Table 3 shows occurrences of Heywood cases out of 500 simulation samples in each of the 16 conditions.

Among many factors contributing to the frequencies of Heywood cases, sample size seems to play an important role. The larger the sample size, the less likely Heywood cases occur. Identifying factors contributing to occurrences of Heywood cases is not the focus of this article, however. The issue has been studied by Anderson and Gerbing (1984) and Van Driel (1978). When Heywood cases occur in bootstrap samples, the decision of whether to include them in construction of bootstrap CIs should be made carefully. Excluding Heywood cases in the bootstrap gives *SE* estimates based on the proportion of samples that contain no Heywood cases; including Heywood cases in the bootstrap gives *SE* estimates based on all samples. Because we are interested in *SEs* of all samples, we include all bootstrap samples regardless of Heywood cases.

Coverage performance. The coverage performance of the bootstrap CIs of θ in each condition is assessed using the empirical coverage rate c,

$$c = \frac{\text{Numbers of CIs including } \theta}{500} \times 100\%.$$

In "model error" conditions, θ are pseudoparameters obtained by fitting the factor analysis model to the population manifest variable correlation matrix. The pseudoparameters in "model error" conditions and parameters in "no model error" conditions are the same, however. In practice a CI can fail to cover the parameter or pseudoparameter θ. If the lower end of the CI is larger than θ, it is referred to as "miss left"; if the upper end of the CI is less than θ, it is referred to as "miss right." The "miss left" rate is the proportion of occurrences of "miss left" in $k = 500$ simulation samples; the "miss right" rate is the proportion of occurrences of "miss right" in $k = 500$ simulation samples. The sum of the "miss left" rate and the "miss right" rate is referred to as the "miss" rate. Figure 3 displays the "miss left," "miss right," and "miss" rates of the five types of bootstrap CIs in the condition where the population manifest variable correlation matrix contains model error, manifest variables are normally distributed, and the sample size is $n = 100$. The "miss" rates of most parameters are around the nominal level of 0.10. The "miss" rates of most parameters are distributed evenly between "miss left" and "miss right." In particular, the three more sophisticated CIs—the BC, the BCa, and the hybrid interval—did not provide much improvement over the *SE*-based bootstrap CI and the percentile interval.

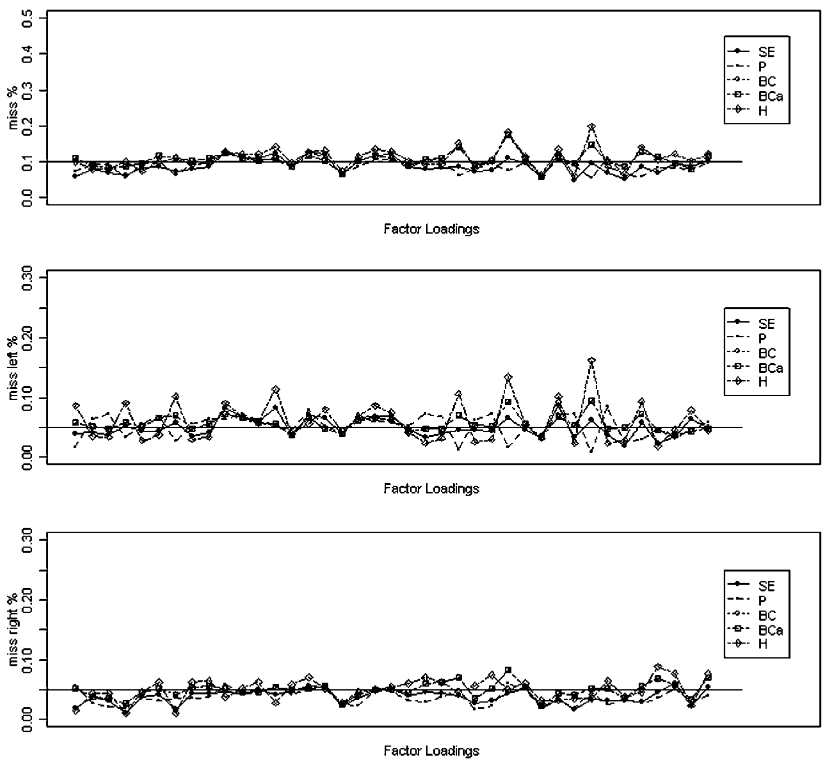

FIGURE 3 Comparison of the five types of bootstrap CIs for factor loadings. The CIs were constructed for the model error condition where the manifest variables are normally distributed and the sample size is $n = 100$. *SE*, P, BC, BCa, and H stand for *SE*-based bootstrap intervals, percentile intervals, bias-corrected percentile intervals, bias-corrected accelerated intervals, and hybrid intervals, respectively. This figure shows the proportion of "miss left," "miss right," and "miss" for the five types of bootstrap CIs.

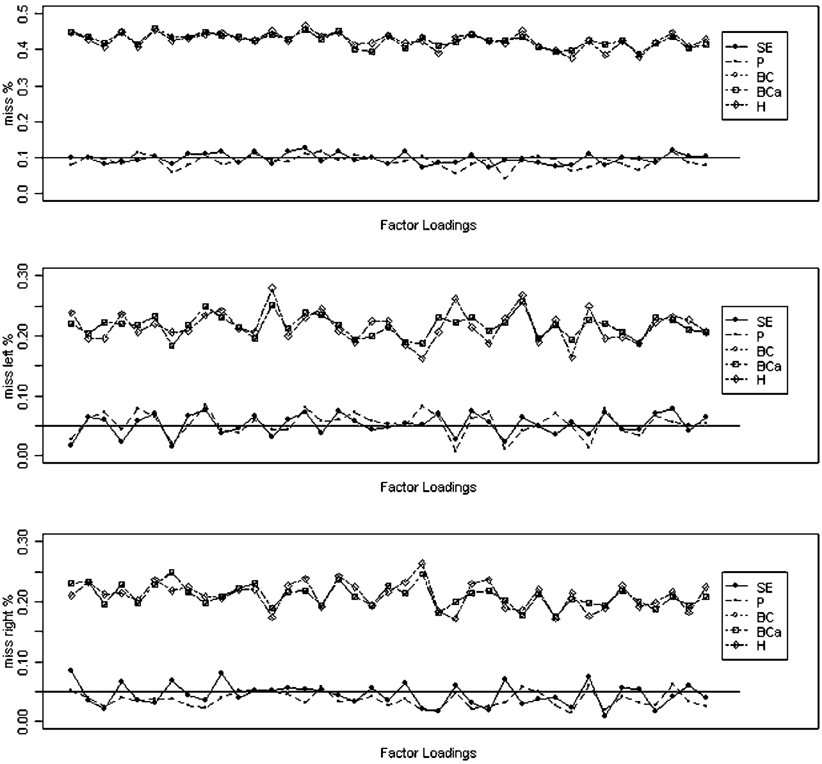

FIGURE 4 Comparison of the five types of bootstrap CIs for factor loadings. The CIs were constructed for the model error condition where the manifest variables are elliptically distributed and the sample size is $n = 100$. *SE*, P, BC, BCa, and H stand for *SE*-based bootstrap intervals, percentile intervals, bias-corrected percentile intervals, bias-corrected accelerated intervals, and hybrid intervals, respectively. This figure shows the proportion of "miss left," "miss right," and "miss" for the five types of bootstrap CIs.

The "miss" rates of most other conditions are similar to this one.[7] However, the more sophisticated CIs performed miserably in the condition where the population manifest variable correlation matrix contains model error, the manifest variable distribution is elliptical, and the sample size is $n = 100$. Figure 4 displays the "miss left," "miss right," and "miss" rates of the five types of bootstrap CIs in this condition. The "miss" rates of the *SE*-based bootstrap

[7]"Miss left %," "miss right %," and "miss %" in all conditions can be found in the supporting file.

intervals and the percentile intervals are around 10%, but "miss" rates of the three more sophisticated intervals are around 40%.

Given the apparently little improvement of the three more sophisticated CIs over the *SE*-based bootstrap intervals and the percentile intervals, we focus on the results of these two simple bootstrap CIs. Because the "miss" rates of the two simple bootstrap CIs do not seem to be predominately "miss left" or "miss right," we report the overall coverage rate, which is one minus the "miss" rate. The empirical coverage rates of the *SE*-based bootstrap intervals and the percentile intervals are presented in Figures 5 (with "model error") and 6 (without "model error"). Each of the two figures includes two panels: the top panel displays normal distribution conditions and the bottom panel displays elliptical distribution conditions. Each box-plot displays the empirical coverage rates of bootstrap CIs of factor loadings at a certain sample size: each

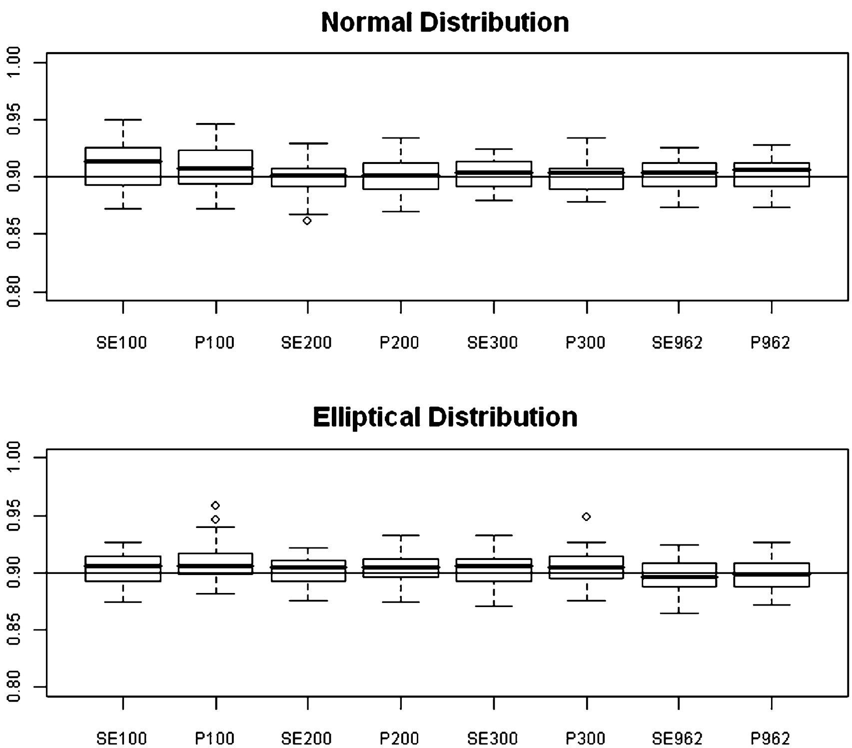

FIGURE 5 Comparison of *SE*-based CIs and percentile intervals for factor loadings in "model error" conditions. *SE* = *SE*-based CIs; P = percentile intervals. The numbers after *SE* or P indicate the sample size.

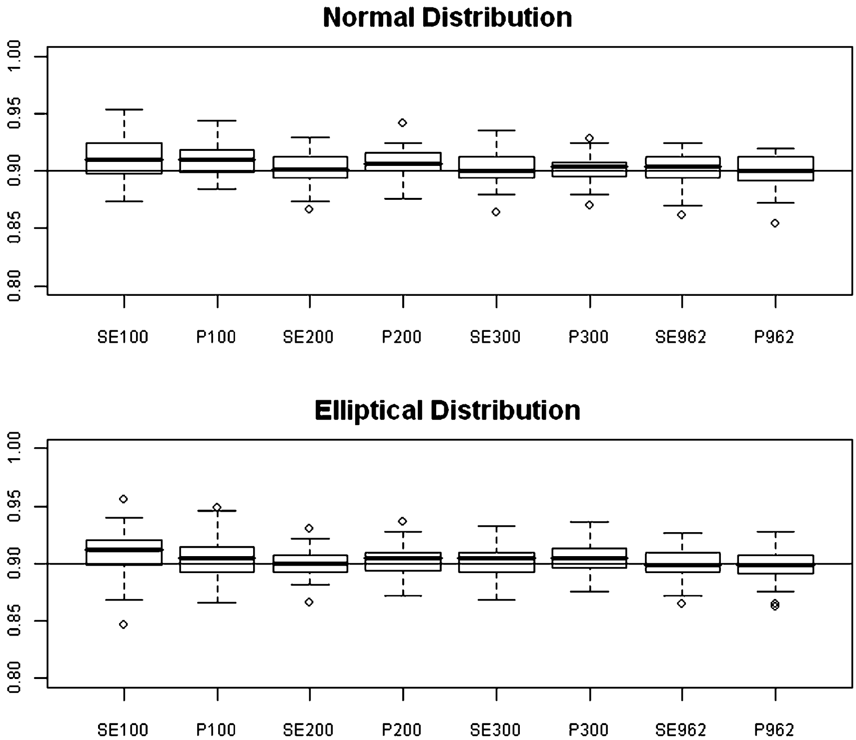

FIGURE 6 Comparison of *SE*-based CIs and percentile intervals for factor loadings in "no model error" conditions. *SE* = *SE*-based CIs; P = percentile intervals. The numbers after *SE* or P indicate the sample size.

point represents the empirical coverage rate of the bootstrap CI for a single factor loading. The empirical coverage rates of the *SE*-based bootstrap CIs and the empirical coverage rates of the percentile intervals from the same set of simulation samples are displayed side by side.

Three observations can be made about the empirical coverage rates of the bootstrap CIs. First, the empirical coverage rates of both types of bootstrap CIs of most parameters are close to the nominal level of 90% for both the normal distribution and the elliptical distribution, for populations with and without "model error," at all sample sizes. The median coverage rates are exactly 90% in many cases. In particular, the bootstrap CIs show good coverage performance when the model is imperfect and the distribution is elliptical. Second, the empirical coverage rates at $n = 100$ are less satisfactory than those of larger sample sizes. Third, the *SE*-based bootstrap CIs and the percentile intervals have comparable empirical coverage rates.

Factor correlations. Factors are correlated in oblique rotation. Table 4 shows empirical coverage rates of the correlation between "limb" and "ear." Because the empirical coverage rates were estimated from 500 simulation samples, we did not expect that they would be exactly 90%. We can test whether the true coverage rate is 90% in a certain condition using a binomial distribution $(500, 0.90)$. Let c be the empirical coverage rate. If $87.4\% < c < 92.6\%$, the decision is not to reject H_0 that the true coverage rate is 90% at the α level of 0.05; otherwise, the decision is to reject H_0. As shown in Table 4, most empirical coverage rates lead to not rejecting H_0 at moderately large sample sizes, for example, $n = 200$. In many cases, the empirical coverage rates are very close to 90%. The empirical coverage rates are lower than 90% at $n = 100$. In particular, the empirical coverage rates of BC, BCa, and hybrid intervals are unsatisfactory at $n = 100$ when the model is imperfect and the distribution is the elliptical distribution.

Figure 7 displays the *SE*-based bootstrap intervals and percentile intervals of the correlation between "limb" and "ear" in the condition combining "model

TABLE 4
Empirical Coverage Rates of Bootstrap CIs for a Factor Correlation in the Simulation Study

Simulation conditions	*n*	*SE-CI*	*Percentile*	*BC*	*BCa*	*Hybrid*
Normal, no model error	100	0.878	0.884	0.856	0.856	0.858
	200	0.874	0.878	0.858	0.858	0.858
	300	0.890	0.894	0.886	0.886	0.884
	962	0.898	0.894	0.896	0.896	0.894
Normal, model error	100	0.884	0.898	0.858	0.858	0.856
	200	0.890	0.890	0.874	0.876	0.882
	300	0.884	0.892	0.878	0.878	0.886
	962	0.900	0.902	0.896	0.896	0.896
Elliptical, no model error	100	0.858	0.886	0.842	0.840	0.844
	200	0.888	0.894	0.874	0.874	0.872
	300	0.912	0.920	0.910	0.910	0.912
	962	0.896	0.898	0.892	0.892	0.904
Elliptical, model error	100	0.890	0.882	0.542	0.542	0.534
Elliptical,	200	0.892	0.894	0.878	0.876	0.874
model error	300	0.918	0.920	0.912	0.912	0.910
	962	0.898	0.902	0.892	0.892	0.896

Note. The parameter of interest is the correlation between factors "limb" and "ear."

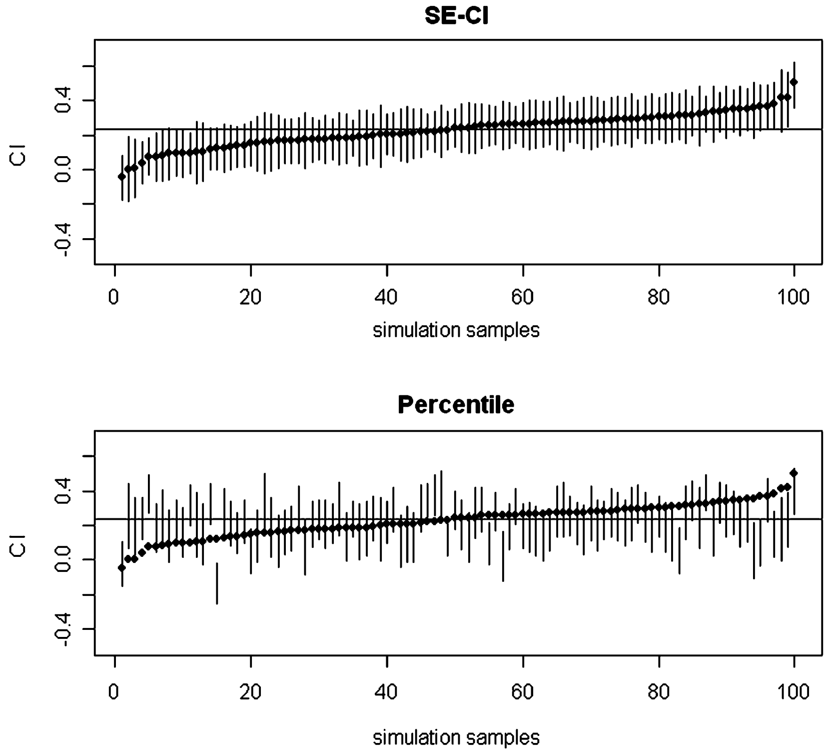

FIGURE 7 *SE*-based CIs and percentile CIs of a factor correlation constructed in the first 100 simulated samples of a "model error" condition. Manifest variables are elliptically distributed and the sample size is $n = 100$. The simulation samples were reordered according to the point estimates. The horizontal line represents the population value of 0.234.

error" and elliptical data. The bootstrap CIs were constructed from the first 100 simulation samples. Though these two types of CIs have comparable empirical coverage rates, their shapes are different. The point estimates are around the center of the *SE*-based bootstrap intervals. The point estimates are at any location relative to the percentile intervals. In some cases, the point estimates are outside of the percentile intervals, which is unpleasant. Note that the point estimates may not be at the very center of the *SE*-based bootstrap intervals because the *SE*-based intervals were constructed using the Fisher Z transformed scale and then transformed back to correlations.

The relations between the point estimates and the percentile intervals suggest the reason the BC and the BCa perform poorly in this condition. When a point estimate is very far away from the median of bootstrap replications, the adjustment with respect to median bias is too severe. In extreme cases, the BC and the BCa can shrink to a single point.

DISCUSSION

The simulation studies showed that all five types of bootstrap CIs had reasonable empirical coverage rates for OLS rotated factor loadings and factor correlations at moderately large sample sizes both with and without model error and for both normal and elliptical distributions. The coverage performances of BC, BCa, and the hybrid interval were not more satisfactory than those of the *SE*-based bootstrap interval and the percentile interval. At the small sample size of $n = 100$, the more sophisticated CIs perform less satisfactorily than the *SE*-based bootstrap interval and the percentile interval.

The results regarding coverage performance accord with a simulation study by Ichikawa and Konishi (2008) on bootstrap CIs of communalities showing that *SE*-based bootstrap intervals and several second-order accurate CIs (Cornish-Fisher, ABC, bootstrap-t, BCa) had comparable coverage performances. They found that "miss" rates of the *SE*-based bootstrap CI tended to concentrate on one side, which we did not replicate in this study. One possible reason is that we considered bootstrap CIs for factor loadings and they considered bootstrap CIs for communalities. Another difference between the simulation studies reported in this article and their simulation study is that we included Heywood cases in constructing bootstrap CIs and they excluded Heywood cases.

Though bias and acceleration correction should produce more accurate CIs in theory (Efron & Tibshirani, 1993, pp. 325–328) the advantages of making such adjustments may require a sample size beyond the reach of a typical study in psychology. Estimating extra parameters like bias and the acceleration constant may reduce bias but will increase variance. The benefits of bias reduction may be offset by the increase in variance at small sample sizes. This is consistent with Yung & Bentler's (1996, p. 223) cautionary note that the bootstrap of structural equation models (of which EFA is a special case) requires a moderately large sample size. Based on an extensive simulation study on orthogonal rotation of maximum likelihood estimates of factor loadings, Ichikawa and Konishi (1995) found that the bootstrap produced satisfactory *SE* estimates for rotated factor loadings. They also compared critical points of the bootstrap-t interval and the hybrid interval with the "actual" critical points. In this study, we directly compared the coverage performance of five types of bootstrap CIs for rotated factor loadings. In addition, our simulation studies were concerned with oblique rotation in both "model error" and "no model error" conditions.

Though OLS is a popular method for estimating factor analysis models, *SE* estimates of OLS factor loadings were unavailable until recently. Browne et al.

(2008) provided analytical *SEs* and CIs for OLS estimates of factor analysis under the condition that manifest variables are normally distributed and model error is not large. When these assumptions are satisfied, the analytical *SE* estimates are a reasonable choice because they require far less computation than the bootstrap method. The bootstrap method provides a convenient tool for constructing CIs when these assumptions appear untenable. For example, MacCallum (2003) pointed out that the factor analysis model is just an approximation even at the population level because the influence of common factors on manifest variables may be nonlinear and many minor common factors are ignored. Yuan and Hayashi (2006, Sec. 3) proved that bootstrap *SE* estimates are consistent in covariance structure modeling, of which the EFA model is a special case. In conditions studied here, bootstrap CIs were accurate for both normal and elliptical data, both with and without model error, and at moderately large sample sizes.

We recommend constructing CIs for rotated factor loadings and factor correlations using bootstrap *SEs* and normal distribution percentiles in a typical study in psychology. Proper transformations of factor correlations, communalities, and unique variances (Browne, 1982) will make the *SE*-based bootstrap CIs lie inside their natural bounds. Other benefits of the *SE*-based bootstrap CIs include easy implementation and relatively light demands for computing.

AN ILLUSTRATION WITH AN EMPIRICAL EXAMPLE

Luo et al. (2008) reported a study on predictors of marital satisfaction using 537 newlywed Chinese couples. In this illustration, we consider an EFA model for the husbands' scores on the revised version of the Chinese Personality Assessment Inventory (CPAI; Cheung et al., 1996). The CPAI includes 28 personality facet scales: each facet scale score is the sum of 10 to 18 items.

We factor analyzed the 28 facet scales using OLS and oblique CF-varimax rotation (Browne, 2001). The RMSEA for the four-factor model is 0.051 with 90% CI $(0.046, 0.056)$. In addition to the five types of bootstrap CIs, we constructed asymptotic CIs of rotated factor loadings and factor correlations under OLS estimation using CEFA 3.02 (Browne et al., 2008). When manifest variables are normally distributed and the model fits the data reasonably well, the asymptotic CIs are preferred because they require far less computation. The facet scale scores do not seem normal, however, because the normalized multivariate kurtosis is 7.17. Bootstrap CIs were constructed with the number of bootstrap sample $B = 2{,}000$.

Table 5 shows the factor loading matrix. Several manifest variables load on more than one factor. For example, "enterprise" loads on "social potency," "dependability," and "relatedness"; "discipline" loads on both "accommodation"

TABLE 5
The Factor Loading Matrix of Luo et al.'s (2008) Survey Data

	Social Potency	*Dependability*	*Accommodation*	*Relatedness*
Novelty	**0.65**	0.03	0.10	0.12
Diversity	**0.50**	−0.15	**0.23**	**0.39**
Divergent thinking	**0.44**	−0.04	−0.03	0.28
Leadership	**0.66**	0.04	**− 0.36**	0.01
Logical-affective	**0.34**	0.10	0.07	**0.40**
Aesthetics	**0.35**	**−0.25**	−0.11	0.24
Extroversion	**0.58**	0.06	−0.01	−0.13
Enterprise	**0.52**	**0.45**	0.12	**− 0.21**
Responsibility	0.05	**0.56**	−0.10	**0.25**
Emotionality	0.01	**− 0.78**	0.01	0.03
Inferiority	**− 0.20**	**− 0.51**	**− 0.38**	−0.12
Practical	−0.05	**0.55**	0.07	**0.31**
Optimism	**0.21**	**0.58**	**0.16**	0.00
Meticulous	−0.12	**0.38**	−0.10	**0.33**
Face	0.07	**− 0.24**	**− 0.30**	0.12
Internal locus	0.09	0.18	**0.42**	−0.10
Family-oriented	0.08	**0.40**	0.24	**0.26**
Defensiveness	0.08	**− 0.20**	**− 0.73**	−0.07
Graciousness	−0.01	**0.40**	**0.53**	0.05
Tolerance	**0.22**	**0.17**	**0.53**	0.09
Self-social	0.07	−0.10	**− 0.60**	0.12
Veraciousness	−0.07	0.17	**0.53**	**0.28**
Traditionalism	−0.05	0.11	**− 0.54**	0.08
Relationship	0.15	0.04	−0.02	**0.61**
Sensitivity	**0.38**	−0.02	−0.17	**0.53**
Discipline	0.00	0.17	**− 0.58**	**0.48**
Harmony	0.00	**0.24**	**0.31**	**0.56**
Thrift	−0.22	−0.02	−0.14	**0.47**

Note. Factor loadings in bold font are those with a 99% *SE*-based bootstrap CI excluding zero. The 99% level is considered because the number of factor loadings is large in the illustration. This is similar to using the Bonferroni correction to control familywise Type I error (Cudeck & O'Dell, 1994).

and "relatedness." Thus the factor pattern demonstrates simple structure but not an independent cluster solution.

Figure 8 displays 90% CIs of the six-factor correlations. Three observations can be made about the plots. First, all five bootstrap CIs are wider than the asymptotic CIs in all six correlations. The difference is due to the large multivariate kurtosis and imperfect model fit. Second, widths of CIs are different for different factor correlations. This highlights the importance of constructing CIs: the point estimates of ρ_{14} and ρ_{24} are similar but their CIs differ substantially. Third, point estimates are around the center of the *SE*-based bootstrap CIs but

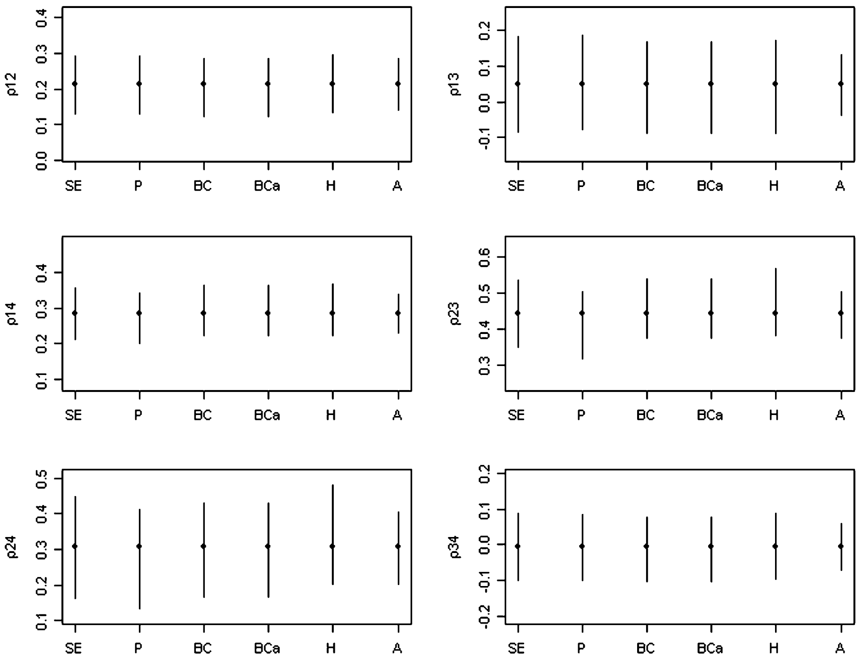

FIGURE 8 Bootstrap CIs for factor correlations of Luo et al.'s (2008) survey data. *SE* = *SE*-based bootstrap CIs; P = percentile intervals; BC = bias-corrected percentile intervals; BCa = bias-corrected accelerated percentile intervals; A = CIs constructed using asymptotic standard errors.

they may be far away from the center of the other four types of bootstrap CIs (as shown in ρ_{23}). Note that point estimates are not necessarily at the very center of the *SE*-based bootstrap CIs because the intervals were constructed using Fisher Z scores and transformed back to correlations. The widths of all five bootstrap CIs are comparable, however.

CONCLUDING COMMENTS

We employed the bootstrap to construct CIs for rotated factor loadings and factor correlations under OLS estimation. A key issue of the bootstrap with EFA is how to deal with the alignment problem. We employed a simulation study to test the column reflection and interchange algorithm and suggested a diagnostic tool for alignment failures. We explored the performance of the *SE*-based bootstrap interval, the percentile interval, the BC, the BCa, and the hybrid interval in a variety of situations. We also illustrated the five types of bootstrap intervals using an empirical study. The results show that the bootstrap CIs have reasonable coverage rates at moderate sample sizes.

The bootstrap is becoming increasingly feasible with the wide availability of inexpensive computing power. For example, it took about 2 min for a 2.66G Dual Core PC to construct bootstrap CIs for the marital satisfaction data. The illustration is nontrivial: it has 28 manifest variables and four factors, and the number of bootstrap samples is $B = 2{,}000$.

Though the bootstrap procedure is extremely versatile and the simulation studies suggest that it works satisfactorily in a number of situations, its use is difficult in the following two situations: First, the column interchange or reflection algorithm might be difficult if the number of factors is large. As computing resources are becoming readily available, intensive computing should be of less concern. Second, certain rotation criteria tend to have multiple solutions, for example, Geomin rotation (Browne, 2001; Browne et al., 2008). Use of the bootstrap procedure for Geomin rotation is not recommended.

An advantage of the bootstrap procedure over the analytical procedure in EFA is that the bootstrap procedure provides valid CIs when the normality assumption is untenable. Our recommendation for factor analysts is to routinely check the normality assumption. If the normality assumption appears appropriate, the analytical procedure is a reasonable choice because of its minimal demand for computing resources. If the normality assumption appears inappropriate, the bootstrap procedure should be considered because of its robustness.

Another advantage of the bootstrap procedure is relatively easy implementation. The same bootstrap procedure can be applied to different estimation methods and rotation methods with little modification. It is particularly useful when new rotation methods are considered.

ACKNOWLEDGMENTS

We are grateful to Michael Browne for numerous helpful conversations and for providing CEFA subroutines. We thank Michael Edwards for reading and commenting on an earlier draft of the article.

REFERENCES

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*, 155–173.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*, 791–799.

Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, *38*, 25–56.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). Cambridge, UK: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*, 62–83.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*, 111–150.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newburn Park, CA: Sage.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2008). *CEFA: Comprehensive Exploratory Factor Analysis.* Retrieved December 2008, from http://faculty.psy.ohio-state.edu/browne/software.php

Chatterjee, S. (1984). Variance estimation in factor analysis: An application of the bootstrap. *British Journal of Mathematical and Statistical Psychology*, *37*, 252–262.

Cheung, F. M., Leung, K., Fan, R., Song, W., Zhang, J., & Zhang, J. (1996). Development of the Chinese Personality Assessment Inventory (CPAI). *Journal of Cross-Cultural Psychology*, *27*, 181–199.

Clarkson, D. B. (1979). Estimating the standard errors of rotated factor loadings by jackknifing. *Psychometrika*, *44*, 297–314.

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

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, *105*, 317–327.

Cudeck, R., & MacCallum, R. C. (2007). *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Erlbaum.

Cudeck, R., & O'Dell, L. L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*, *115*, 475–487.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

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.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Mahwah, NJ: Erlbaum.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *The Annals of Statistics*, *16*, 927–953.

Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. *Psychometrika*, *60*, 77–93.

Ichikawa, M., & Konishi, S. (2008). Constructing second-order accurate confidence intervals for communalities in factor analysis. *British Journal of Mathematical and Statistical Psychology*, *61*, 361–378.

Jennrich, R. I. (1973a). On the stability of rotated factor loadings: The Wexler phenomenon. *British Journal of Mathematical and Statistical Psychology*, *26*, 167–176.

Jennrich, R. I. (1973b). Standard errors for obliquely rotated factor loadings. *Psychometrika*, *38*, 593–604.

Jennrich, R. I. (2007). Rotation methods, algorithms, and standard errors. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 315–335). Mahwah, NJ: Erlbaum.

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

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1991). Approximating confidence intervals for factor loadings. *Multivariate Behavioral Research*, *26*, 421–434.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworth & Company, Ltd.

Luo, S., Chen, H., Yue, G., Zhang, G., Zhaoyang, R., & Xu, D. (2008). Predicting marital satisfaction from self, partner, and couple characteristics: Is it me, you, or us? *Journal of Personality*, *76*, 1231–1266.

MacCallum, R. C. (2003). Working with imperfect models. *Multivariate Behavioral Research*, *38*, 113–139.

MacCallum, R. C., Browne, M. W., & Cai, L. (2007). Factor analysis models as approximations. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 153–175). Mahwah, NJ: Erlbaum.

MacCallum, R. C., Tucker, L. R., & Briggs, N. E. (2001). An alternative perspective on parameter estimation in factor analysis and related mothods. In R. Cudeck & S. Du Toit (Eds.), *Structural equation modeling: Present and future* (pp. 39–57). Chicago: Scientific Software International Inc.

Pennel, R. (1972). Routinely computable confidence intervals for factor loadings using the "jack-knife." *British Journal of Mathematical and Statistical Psychology*, *25*, 107–114.

Porac, C., Coren, S., Steiger, J. H., & Duncan, P. (1980). Human laterality: A multidimensional approach. *Canadian Journal of Psychology*, *34*, 91–96.

Thurstone, L. L. (1947). Multiple factor analysis. Chicago: University of Chicago Press.

Van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika*, *43*, 225–243.

Woods, C. M., & Edwards, M. C. (2008). Factor analysis and related methods. In C. R. Rao, J. P. Miller, & D. C. Rao (Eds.), *Handbook of statistics 27: Epidemiology and medical statistics* (pp. 367–394). New York: Elsevier.

Yanai, H., & Ichikawa, M. (2007). Factor analysis. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics 26: Psychometrics* (pp. 257–296). New York: Elsevier.

Yuan, K., & Hayashi, K. (2006). Standard errors in covariance structure models: asymptotics versus bootstrap. *British Journal of Mathematical and Statistical Psychology*, *59*, 397–417.

Yuan, K., Marshall, L. L., & Bentler, P. M. (2002). A unified approach to exploratory factor analysis with missing data, nonormal data, and in the presence of outliers. *Psychometrika*, *67*, 95–122.

Yung, Y.-F., & Bentler, P. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 195–226). Mahwah, NJ: Erlbaum.