American Psychological Association

Psychological Methods

1082-989X/19/$12.00

http://dx.doi.org/10.1037/met0000230

Determining the Number of Factors Using Parallel Analysis and Its Recent Variants

Sangdon Lim and Seungmin Jahng
Sungkyunkwan University

Abstract

Parallel analysis (PA) is recommended as one of the best procedures to determine the number of factors but its theoretical justification has long been questioned. The current study discussed theoretical issues on the use of eigenvalues for dimensionality assessment and reviewed the development of PA and its recent variants proposed to address the issues. The performances of 13 different PAs including PA with minimum rank factor analysis, revised PA, and comparison data method were investigated through a Monte Carlo simulation under a wide range of factor structures that produce small factor-representing and nonrepresenting eigenvalues for different types of measurement scales. Results showed that the traditional PA using full correlation matrices performed best in most of the conditions, especially when population error was involved. However, the overall accuracy of PA was not satisfactory when factor-representing eigenvalues were small, that is, when factor loadings were low and factor correlations were high. From these results, we suggest that the original PA be used to determine the number of factors but the estimated number should not be taken as a fixed estimate. The number of factors within ± 1 range of the estimate can be considered as viable candidates and interpretational validity of the competing models should be consulted for the decision.

Translational Abstract

Determining the number of factors is one of the most important decisions in exploratory factor analysis in psychological studies. Parallel analysis (PA) which compares eigenvalues of the sample data with those of random data is one of the most recommended procedures for the decision by many experts. However, there are unsolved theoretical issues in PA and alternative procedures to address the issues have been proposed in the literature. The current study discussed several issues on the use of eigenvalues for determining the number of factors and reviewed the development of PA and its alternatives. We also examined the performances of 13 PA variants using a comprehensive simulation study. We found that the original PA was the most accurate in most conditions but its overall accuracy was not satisfactory when factors are highly correlated or factor-variable relations are not strong enough. In summary, we suggest that PA and its alternatives should be used with caution and the interpretability of competing factor models should be consulted to determine the number of factors.

Keywords: parallel analysis, revised parallel analysis, comparison data method, minimum rank factor analysis, number of factors

One of the biggest challenges in exploratory factor analysis (EFA) is determining the number of common factors underlying a set of variables (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Fava & Velicer, 1992). Researchers proposed several procedures estimating the number of factors to retain (e.g., the Kaiser rule, chi-square test for residual correlations, and parallel analysis) and evaluated the accuracy of the procedures using simulation studies (e.g., Humphreys & Montanelli, 1975; Preacher, Zhang, Kim, & Mels, 2013; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). Among the various procedures, parallel analysis (Horn, 1965) has been recommended as one of the best methods by many experts (e.g., Fabrigar et al., 1999; Floyd & Widaman, 1995; Hayton, Allen, & Scarpello, 2004; Preacher & MacCallum, 2003).

Parallel analysis (PA) compares the eigenvalues of the sample correlation matrix with the eigenvalues obtained from a random correlation matrix for which no factors are assumed. The original PA procedure (Horn, 1965) is known to be relatively accurate, but its theoretical justification has long been questioned. First, the eigenvalues of the reduced correlation matrix with communalities

Sangdon Lim and Seungmin Jahng, Department of Psychology, Sungkyunkwan University.

Sangdon Lim is now at Department of Educational Psychology, University of Texas at Austin.

Part of the data and the narratives have been previously disseminated at the semi-annual meeting of Korean Society of Psychological Measurement and Assessment held in Seoul, South Korea in May 2017.

Correspondence concerning this article should be addressed to Seungmin Jahng, Department of Psychology, Sungkyunkwan University, 25-2 Sungkyunkwan-ro, Jongno-gu, Seoul 03063, South Korea. E-mail: jahngs@skku.edu

of the variables in the main diagonal have a direct relation to the number of factors, but those of the full correlation matrix with ones in the diagonal do not (Guttman, 1954; Mulaik, 2010; Timmerman & Lorenzo-Seva, 2011). Second, the comparison of the sample eigenvalues with the random eigenvalues from the zero-factor population is justified for the first eigenvalue but not for the rest (Braeken & van Assen, 2017; Green, Levy, Thompson, Lu, & Lo, 2012; Ruscio & Roche, 2012; Turner, 1998).

To address the first limitation, Humphreys and Ilgen (1969) introduced a modified PA procedure that uses a reduced correlation matrix where the squared multiple correlation (SMC) of each variable is used as a communality estimate. Recently, Timmerman and Lorenzo-Seva (2011) proposed using minimum rank factor analysis (MRFA; ten Berge & Kiers, 1991) in PA to obtain better communality estimates. For the second problem, Green, Levy, Thompson, Lu, and Lo (2012) proposed the revised parallel analysis (RPA) that updates the population model to a series of structures with nonzero factors from which random eigenvalues are generated. Ruscio and Roche (2012) also developed a similar procedure called comparison data method (CD) that updates the population model to a series of bootstrapped structures with nonzero factors. Several studies evaluated the performance of the alternative PA procedures and recommended the use of the alternatives rather than the original PA in many conditions (Crawford et al., 2010; Green et al., 2012; Ruscio & Roche, 2012; Timmerman & Lorenzo-Seva, 2011).

In the previous studies, the theoretical problems of the original PA were discussed and addressed in part, but a more general discussion in a broader context was limited. In addition, the factor structures and other conditions used to evaluate the performance of the procedures are either incomprehensive in each study or inconsistent across the studies, making it difficult to fully understand the nature of the procedures and to identify the source of the difference in performance. In this study, we reviewed the theoretical issues in PA and the development of PA and its variants that address those issues, in the context of using eigenvalues for dimensionality assessment. We also evaluated the performance of the PA and its variants and compared their accuracy across a wide range of factor structures and data conditions to better understand how the suggested procedures work.

Dimensionality Assessment Using Eigenvalues and Parallel Analysis

The most commonly used information when determining the number of factors is the eigenvalues of the sample correlation matrix. Let $\mathbf{\Sigma}$ be a $p \times p$ population correlation matrix. A k-factor model for $\mathbf{\Sigma}$ can be written as

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}} + \mathbf{\Psi}, \quad (1)$$

where $\mathbf{\Lambda}$ is a $p \times k$ pattern matrix, $\mathbf{\Lambda}^{\mathbf{T}}$ is the transpose of $\mathbf{\Lambda}$, $\mathbf{\Phi}$ is a $k \times k$ factor correlation matrix, and $\mathbf{\Psi}$ is a $p \times p$ diagonal uniqueness matrix. The fundamental theoretical basis for using eigenvalues to determine the number of factors k is that k is equal to the rank of $\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$ (Mulaik, 2010). The $p \times p$ matrix $\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$ has p eigenvalues, of which k eigenvalues are positive and the remaining $p - k$ eigenvalues are equal to zero. Such a matrix that has only non-negative eigenvalues is called a Gramian matrix. Guttman (1954) showed that the number of eigenvalues of $\mathbf{\Sigma}$ equal to or greater than 1.0 is one lower bound for the minimum rank of $\mathbf{\Sigma} - \mathbf{\Psi} = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$. However, this does not hold in the sample correlation matrix $\mathbf{S}$ because $\mathbf{S} - \mathbf{\Psi} \neq \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$, and $\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$ and $\mathbf{\Psi}$ are unknown in the sample. For this reason, the use of sample eigenvalues in determining the number of factors lacks mathematical justification and must be regarded as a heuristic (Guttman, 1954; Mulaik, 2010).

Moreover, it is difficult to develop an accurate dimensionality assessment procedure based on eigenvalues because the eigenvalues of the sample correlation matrix are influenced not only by the true factor structure but also by at least three other sources: (a) the least-squares bias in the sample (Horn, 1965); (b) the use of communality estimates (Humphreys & Ilgen, 1969); and (c) the constraint between eigenvalues (Turner, 1998). The development and modification of the parallel analysis have been facilitated by taking each source into account in the eigenvalue-based dimensionality assessment.

Least-Squares Bias in Sample

Dimensionality assessment procedures based on eigenvalues often involve sequential classifications, where each eigenvalue is classified into a signal or noise, with a signal indicating a true factor. Kaiser (1960) suggested that 1.0 could be used as a classification threshold when developing a heuristic procedure based on sample eigenvalues. In the Kaiser rule, if a sample eigenvalue is larger than 1.0 then it is classified as a signal, otherwise it is classified as noise. The number of signals is used as an estimate of k, the number of common factors. However, the Kaiser rule overestimates k because of the least-squares bias in the sample (Horn, 1965): The sample eigenvalues are determined by the sampling variation of the correlations as well as the underlying factor structure. The random variation causes biased estimates of the population eigenvalues, creating a positive bias for the first several sample eigenvalues and a negative bias for the rest. One result of the bias is that the first half of the eigenvalues of the sample correlation matrix are expected to be greater than 1.0, even if the variables are completely uncorrelated in the population.

Horn (1965) proposed a Monte Carlo method to obtain a set of distributions of the reference eigenvalues, conditional to $n \times p$, the size of the sample data. Parallel analysis (Horn, 1965) generates multiple data sets from the zero-factor p-variate normal population to construct a sampling distribution of random eigenvalues for each of the p sample eigenvalues. The original PA used the average of each sampling distribution as a classification threshold for the corresponding sample eigenvalue (PA-PCA-m; see Table 1 for the list of acronyms), but later studies suggested that using the 95th percentile of each distribution as a threshold generally increase the accuracy (PA-PCA-95; Buja & Eyuboglu, 1992; Glorfeld, 1995). The detailed steps of PA are shown in Table 2.

Use of Communality Estimates

Another issue in determining the number of factors based on eigenvalues is related to whether to use and how to estimate the reduced correlation matrix. Factor analysis assumes that the sample data is generated from a common factor model, where the

Table 1
List of the Evaluated Procedures

Acronym	Procedure	Type of correlation matrix	Elements of distribution	Threshold
PA-PCA-m	Parallel analysis	Full	Eigenvalues	Mean
PA-PCA-95	Parallel analysis	Full	Eigenvalues	95th percentile
PA-PAF-m	Parallel analysis	Reduced (SMC)	Eigenvalues	Mean
PA-PAF-95	Parallel analysis	Reduced (SMC)	Eigenvalues	95th percentile
PA-MRFA-eg-m	Parallel analysis	Reduced (MRFA)	Eigenvalues	Mean
PA-MRFA-eg-95	Parallel analysis	Reduced (MRFA)	Eigenvalues	95th percentile
PA-MRFA-m	Parallel analysis	Reduced (MRFA)	ECV values	Mean
PA-MRFA-95	Parallel analysis	Reduced (MRFA)	ECV values	95th percentile
RPA-PCA-m	Revised parallel analysis	Full	Eigenvalues	Mean
RPA-PCA-95	Revised parallel analysis	Full	Eigenvalues	95th percentile
RPA-PAF-m	Revised parallel analysis	Reduced (SMC)	Eigenvalues	Mean
RPA-PAF-95	Revised parallel analysis	Reduced (SMC)	Eigenvalues	95th percentile
CD	Comparison data method	Full	RMSR values	$\alpha = .30$[a]

Note. SMC = squared multiple correlation; PAF = principal axis factoring; MRFA = minimum rank factor analysis; ECV = explained common variance; RMSR = root mean squared residuals between sample and random eigenvalues.
[a] Significance level for Mann-Whitney *U* test of two RMSR distributions.

variance of each variable is composed of two components: variance inherited from common factors (communality) and variance unique to the variable (uniqueness). As described in Equation 1, the eigenvalues of the reduced correlation matrix ($\boldsymbol{\Sigma} - \boldsymbol{\Psi}$) with communalities in the diagonal indicate the number of factors but Horn's PA uses the eigenvalues from the full correlation matrix $\boldsymbol{\Sigma}$. This is the same as assuming that all variables are devoid of their unique aspects, that is, $\boldsymbol{\Psi} = \mathbf{0}$.

Table 2
Detailed Steps of the Estimation Procedures

Steps	PA	RPA	CD
Step 1 Obtaining sample values	Calculate $\mathbf{S}$ from $\mathbf{X}$. Obtain $\boldsymbol{\lambda}_\mathbf{S}$ from [reduced] $\mathbf{S}$. Store $\boldsymbol{\tau}$ if $\mathbf{X}$ is ordinal.	Calculate $\mathbf{S}$ from $\mathbf{X}$. Obtain $\boldsymbol{\lambda}_\mathbf{S}$ from [reduced] $\mathbf{S}$. Store $\boldsymbol{\tau}$ if $\mathbf{X}$ is ordinal.	Calculate $\mathbf{S}$ from $\mathbf{X}$. Obtain $\boldsymbol{\lambda}_\mathbf{S}$ from $\mathbf{S}$.
Step 2 Initializing	Set $j = 0$.	Set $j = 0$.	Set $j = 1$.
Step 3 Modeling population		If $j = 0$, set $\boldsymbol{\Sigma}_\mathbf{j} = \mathbf{I_p}$. If $j > 0$, fit a j-factor EFA model on $\mathbf{S}$ and obtain $\boldsymbol{\Lambda}_\mathbf{j}$. Calculate $\boldsymbol{\Sigma}_\mathbf{j}^* = \boldsymbol{\Lambda}_\mathbf{j}\boldsymbol{\Lambda}_\mathbf{j}^\mathbf{T}$ and replace the diagonals with 1.0 to obtain $\boldsymbol{\Sigma}_\mathbf{j}$.	Generate a j-factor bootstrapped dataset $\mathbf{X}_\mathbf{j}^*$ with GenData module.
Step 4 Constructing reference distributions	Sample $\mathbf{X_0}$ from $\mathbf{N_p}(\mathbf{0}, \mathbf{I_p})$. Convert to an ordinal dataset with $\boldsymbol{\tau}$ if $\mathbf{X}$ is ordinal. Calculate $\mathbf{R_0}$ from $\mathbf{X_0}$. Obtain and store $\boldsymbol{\lambda}_\mathbf{0}$ from [reduced] $\mathbf{R_0}$. Repeat *Step 4* sufficiently to construct a distribution for each i of $\lambda_{0(i)}$.	Sample $\mathbf{X_j}$ from $\mathbf{N_p}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{j})$. Convert to an ordinal dataset with $\boldsymbol{\tau}$ if $\mathbf{X}$ is ordinal. Calculate $\mathbf{R_j}$ from $\mathbf{X_j}$. Obtain $\boldsymbol{\lambda}_\mathbf{j}$ and store its $\lambda_{j(j+1)}$ from [reduced] $\mathbf{R_j}$. Repeat *Step 4* sufficiently to construct a distribution of $\lambda_{j(j+1)}$.	Sample $\mathbf{X_j}$ from $\mathbf{X}_\mathbf{j}^*$. Calculate $\mathbf{R_j}$ from $\mathbf{X_j}$. Obtain $\boldsymbol{\lambda}_\mathbf{j}$ from $\mathbf{R_j}$. Calculate and store the RMSR between $\boldsymbol{\lambda}_\mathbf{S}$ and $\boldsymbol{\lambda}_\mathbf{j}$. Repeat *Step 4* sufficiently to construct a j-factor distribution of RMSR values.
Step 5 Deciding sufficiency	From the $(j + 1)$-th distribution, take the mean [or 95th percentile] as threshold. If $\lambda_{S(j+1)}$ exceeds the threshold, increment j by 1 and repeat *Step 5*.	From the $(j + 1)$-th distribution, take the mean [or 95th percentile] as threshold. If $\lambda_{S(j+1)}$ exceeds the threshold [and $\lambda_{S(j+1)} > 0$ if $\mathbf{S}$ is reduced], increment j by 1 and repeat from *Step 3*.	Compare a j-factor distribution with a $(j + 1)$-factor distribution using Mann-Whitney *U* test ($\alpha = .30$). If the difference is significant, increment j by 1 and repeat from *Step 3*.
Step 6 Finalizing	Return j as the estimated number of factors k.	Return j as the estimated number of factors k.	Return j as the estimated number of factors k.

Note. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; [text] = steps for the variants of the estimation procedures; $\mathbf{X}$ = sample data ($n \times p$); n = the number of observations; p = the number of variables; $\mathbf{S}$ = sample correlation matrix ($p \times p$); Pearson or polychoric; $\boldsymbol{\tau}$ = thresholds for categorical scales in $\mathbf{X}$, if $\mathbf{S}$ is polychoric; $\mathbf{X_j}$ = j-factor random dataset ($n \times p$), $j = 0, \ldots, p - 1$; k = estimated number of factors; $\mathbf{N_p}$ = p-variate normal distribution; $\mathbf{I_p}$ = identity matrix of order p; $\mathbf{R_j}$ = random correlation matrix with j factors ($p \times p$), Pearson or polychoric; $\boldsymbol{\lambda}_\mathbf{S}$ = a set of p sample eigenvalues or ECVs $\{\lambda_{S(i)}; i = 1, \ldots, p\}$; ECV = explained common variance; $\boldsymbol{\Lambda}_\mathbf{j}$ = unrotated j-factor loading matrix ($p \times j$); $\boldsymbol{\Sigma}_\mathbf{j}$ = j-factor population structure ($j \times j$); $\mathbf{X}_\mathbf{j}^*$ = bootstrapped population (10,000 × p) with a j-factor structure; $\boldsymbol{\lambda}_\mathbf{j}$ = a set of p random eigenvalues or ECVs from a j-factor structure $\{\lambda_{j(i)}; i = 1, \ldots, p\}$; RMSR = root-mean-square residuals between two eigenvalue sets $\boldsymbol{\lambda}_\mathbf{S}$ and $\boldsymbol{\lambda}_\mathbf{j}$, obtained by $\text{RMSR}_j = [\Sigma(\lambda_{S(i)} - \lambda_{j(i)})^2/p]^{1/2}$. Reduced $\mathbf{S}$ or $\mathbf{R}$ is a matrix where each diagonal element is replaced with a communality estimated with SMC or MRFA.

Humphreys and Ilgen (1969) argued that PA should assume the unique variance and suggested using the reduced correlation matrix when estimating the number of common factors. Guttman (1954) showed that the number of positive eigenvalues of $(\boldsymbol{\Sigma} - \boldsymbol{\Psi})$ in Equation 1 is a lower bound for the minimum rank of $(\boldsymbol{\Sigma} - \boldsymbol{\Psi})$, where the SMC of each variable is placed on the diagonal of $(\boldsymbol{\Sigma} - \boldsymbol{\Psi})$. Because communalities are unknown in the sample, Humphreys and Ilgen (1969) suggested using the sample SMC of each variable as an estimate of communality in parallel analysis. In this modified PA procedure, the sample eigenvalues and the random eigenvalues are obtained from the reduced correlation matrix instead of the full correlation matrix. Crawford et al. (2010) refers to this variant as PA with principal axis factoring (PA-PAF), and the original as PA with principal component analysis (PA-PCA; see Table 1).

Timmerman and Lorenzo-Seva (2011) highlighted that a reduced correlation matrix with SMCs as communality estimates often has negative eigenvalues, and hence cannot be reliably interpreted as related to the explained variance of a factor and thus the number of factors. They proposed using minimum rank factor analysis in PA (PA-MRFA). MRFA provides communality estimates under the constraint that the reduced correlation matrix has only non-negative eigenvalues. PA-MRFA compares the explained common variance (ECV) of the MRFA solution (MRFA-ECV) obtained from the sample data with the sampling distribution of MRFA-ECV obtained from random data sets. The MRFA-ECV is calculated by dividing each eigenvalue by the sum of all eigenvalues, which are obtained from the MRFA reduced correlation matrix. However, the distribution of eigenvalues can be used instead of the distribution of ECV in PA-MRFA (Garrido, Abad, & Ponsoda, 2013).

Constraint Between Eigenvalues

For a $p \times p$ correlation matrix, the sum of p eigenvalues is always p. Because of this constraint, the size of the first eigenvalue limits the size of the remaining eigenvalues. Thus, because the first eigenvalue of a $p \times p$ correlation matrix of a one-factor structure is larger than the first eigenvalue of a $p \times p$ correlation matrix of a zero-factor structure, the second eigenvalue of the one-factor structure is smaller than the second eigenvalue of the zero-factor structure. In general, the $(j + 1)$th eigenvalue of a j-factor correlation matrix is smaller than the $(j + 1)$th eigenvalue of a zero-factor correlation matrix, for $j = 1, 2, \ldots, p - 1$. Thus, the sampling distribution (and the threshold) of the second and subsequent eigenvalues obtained in PA is positively biased for a k-factor model where $k > 1$. This leads to an increased false negative error in PA by incorrectly classifying a small signal as noise, underestimating the number of factors. For the second and subsequent factors, the existence of the $(j + 1)$th factor must be evaluated by comparing the $(j + 1)$th sample eigenvalue against the $(j + 1)$th eigenvalues generated from the j-factor structure (Harshman & Reddon, 1983; Saccenti & Timmerman, 2017; Turner, 1998).

Recent studies proposed alternative stepwise procedures that make sequential adjustments by incrementing j in each step. One such procedure is the revised parallel analysis (Green et al., 2012). In RPA, the sampling distribution of the $(j + 1)$th eigenvalue is constructed from multiple sets of random data generated from a j-factor structure. The j-factor structure is estimated from the sample data for $j > 0$. When evaluating the first eigenvalue, that is, for $j = 0$, RPA is performed in the same way as the original PA. For the second and subsequent eigenvalues, that is, $j > 0$, if the $(j + 1)$th sample eigenvalue is greater than the mean or 95th percentile of the $(j + 1)$th eigenvalues generated from the j-factor structure, the hypothesis that the number of factors k is equal to j is rejected. Otherwise, the hypothesis is retained. Panel (a) in Figure 1 illustrates an example of RPA using a sample data set generated from a $k = 2$ structure. In Panel (a) where the second sample eigenvalue is tested, PA-PCA-95 fails to detect the second eigenvalue as a signal but RPA-PCA-95 correctly classifies it.

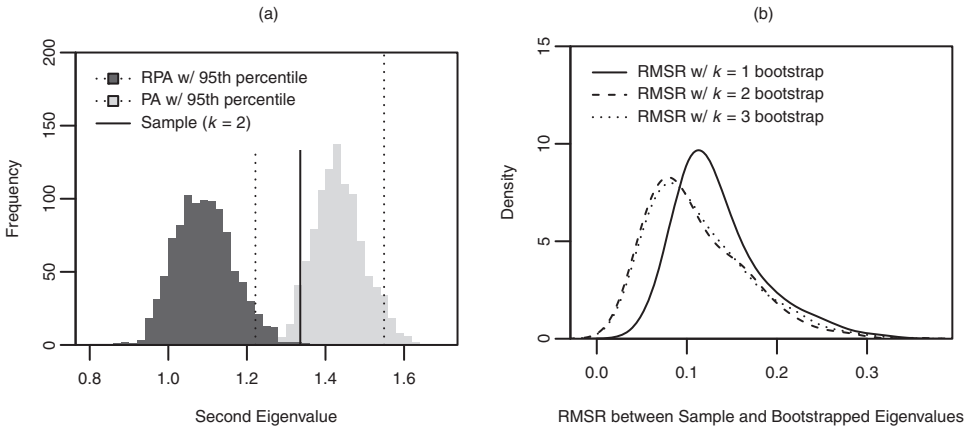

Figure 1. Graphical illustrations of the revised parallel analysis (a) and the comparison data method (b). The sample data set with 100 observations was generated from a two-factor structure with each factor loaded onto eight and four items at .8 and .5, respectively and factor correlation of .5. The second sample eigenvalue in Panel (a) (solid vertical line) is smaller than the 95th percentile of the second PA eigenvalues (dotted vertical line on right), but larger than that of the second RPA eigenvalues (dotted vertical line on left). In Panel (b), the distribution of RMSR for the two-factor model is significantly different from the one-factor model but not from the three-factor model.

Another alternative is the comparison data method (Ruscio & Roche, 2012). Using a fit comparison approach, CD examines whether the sample eigenvalues are better fitted by multiple sets of random data from a j-factor structure or a $(j + 1)$-factor structure. The evaluation of the fit is achieved by constructing the distribution of root mean squared residuals (RMSR) between the sample eigenvalues and random eigenvalues generated from a bootstrapped population with a specified number of factors. The bootstrapped population for a factor structure can be obtained from the sample data using a specialized procedure, called GenData (Ruscio & Kaczetow, 2008). If the RMSR distribution for the $(j + 1)$-factor structure is significantly different from that for the j-factor structure, the hypothesis of $k = j$ is rejected. The significance test can be conducted using Mann–Whitney U test with 500 RMSR values and $\alpha = .30$, for example. Panel (b) in Figure 1 shows an example of CD using the same data set used in Panel (a) in Figure 1. The two-factor model has a better RMSR distribution than the one-factor model but the fit does not improve significantly with the three-factor model. The detailed steps of the two decision procedures are presented in Table 2.

Sources of Small Signal Eigenvalues

Given that the original PA tends to underestimate the number of factors, the very concern that the PA variants have in common is to detect a small jth eigenvalue that indicates the presence of the jth factor, $j = 1, \ldots, p$. We describe two key features of the factor model that produce a small signal eigenvalue that must be detected.

The variance explained by each factor is essentially related to the sum of the squared loadings for the factor. Therefore, if the number of variables related to a factor is small and the factor loadings are low in magnitude, the eigenvalue corresponding to the factor becomes small. Panel (a) in Figure 2 shows the average of the second eigenvalues of 1,000 samples, which depends on the number of variables for the second factor and their factor loadings. Each sample was generated from an orthogonal two-factor structure with 100 observations. The first factor has four variables with the same factor loadings of .4. The average of the second sample eigenvalues decreases as the loadings for the second factor decreases. When the second factor is measured with four variables, the average is less than the PA threshold when the loadings are equal to or lower than .34. With only three variables for the second factor, the average is less than the PA threshold even when the loadings are .40.

The second source of the small signal eigenvalue is related to factor correlation. In principal component analysis or principal axis factoring, each eigenvalue is extracted such that the variance explained by the corresponding component (or axis) does not overlap with that explained by the preceding component (i.e., orthogonal). This property makes the correlation matrix produce a larger first eigenvalue and smaller successive eigenvalues when the underlying factors are strongly correlated. Panel (b) in Figure 2 shows an example of a two-factor structure with varying factor correlations, where each factor has four variables with the same loading of .4. With a nonzero factor correlation, the variance captured by the first axis is not solely due to the first factor. This is because the first axis captures not only the common variance of the variables directly explained by the first factor but also the common variance indirectly explained by the same factor through the second factor correlated with the first factor. Accordingly, the magnitude of the first eigenvalue, as an index of the presence of a factor, is inflated by the factor correlation, leaving a smaller variance that can be captured by the second axis, that is, a smaller second eigenvalue.

These examples illustrate that low factor loadings, a small number of variables, and/or strong factor correlations will produce small signal eigenvalues, which potentially leads to misclassification of signal eigenvalues as noise. Failing to detect a small signal eigenvalue may occur more frequently for the second and following factors than the first factor in PA with the inflated thresholds.

Considerations When Evaluating the Performance of PA Variants

Some of the simulation studies on PA have compared the performance of different PA procedures (Achim, 2017; Crawford et al., 2010; Garrido et al., 2013; Green et al., 2012; Green, Redell, Thompson, & Levy, 2016; Green, Thompson, Levy, & Lo, 2015;

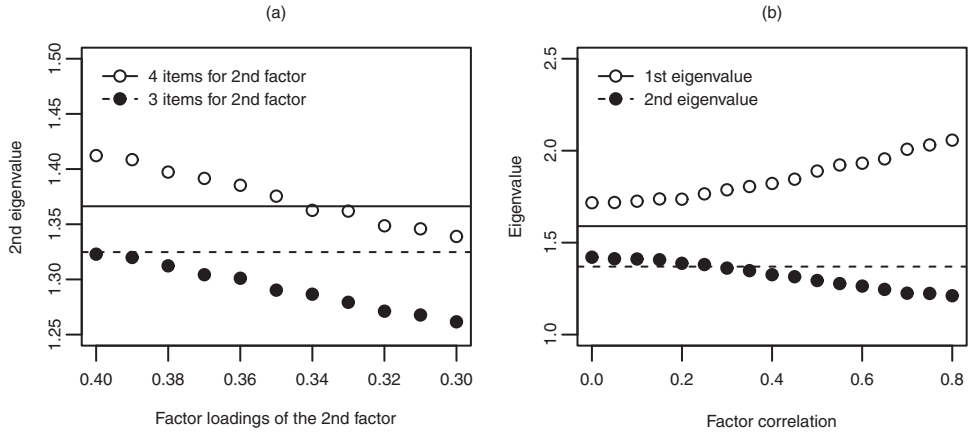

Figure 2. Effects of the number of items and factor loadings (a) and factor correlation (b) on the eigenvalues of a sample correlation matrix of the two-factor model. The lines represent PA-PCA-95 thresholds.

Ruscio & Roche, 2012; Timmerman et al., 2011). However, because the procedures and conditions examined vary from study to study, it is difficult to draw general conclusions about the performance of the procedures. Furthermore, in order to reach conclusions that can be applied to empirical studies using real data, the performance of PA procedures must be investigated under a wide range of factor structures with small signal and noise eigenvalues. The two key features that produce small signal eigenvalues are the high factor correlation and the small sum of squared loadings. A promising PA procedure would perform accurately in the presence of the two features, each or combined. Therefore, it is very important to manipulate both factor correlation and the sum of squared loadings when evaluating the PA variants. As for the sum of the squared loadings, it is useful to investigate the performance by manipulating the magnitude of loadings for the factors measured by four or more variables because it is generally suggested that a factor should be measured by four or more variables (Fabrigar et al., 1999; Floyd & Widaman, 1995).

Demonstrating the capability to detect small signals would be better accompanied by demonstrating the capability to dismiss weak noise. The practical utility of the PA procedures would be better gauged by evaluating their performance when there is a small noise eigenvalue, which can be manipulated by introducing noise factors or population error in the factor structure. The idea behind population error is that a common factor model will never fit exactly at the population level and provide an approximation at best (MacCallum & Tucker, 1991; Preacher et al., 2013). Population error in the common factor model is sometimes referred to as model error (MacCallum & Tucker, 1991), minor factors (Timmerman & Lorenzo-Seva, 2011), or trivial factors (Zwick & Velicer, 1986). Small noise eigenvalues can also be introduced in a factor model if variables are skewed in opposite directions, likely producing spurious factors known as difficulty factors. Difficulty factors derived from skewed variables are not considered as true factors (Gorsuch, 1983; Olsson, 1979; McDonald & Ahlawat, 1974).

Generalizability of the findings is another important consideration. Results from factor structures with a small number of fixed parameter values (e.g., a fixed set of factor loadings and factor correlations) may be less generalizable than results from factor structures with random parameter values. Only a few PA procedures have been assessed for their performance in factor structures with random parameters (e.g., Ruscio & Roche, 2012). Evaluation of the proposed PA procedures with such realistic factor structures would be of practical importance. The PA procedures can also be better assessed by examining their performance on variables with different types of measurement scales. For continuous variables, PA is applied to the Pearson correlation matrix. For ordered categorical variables, it is generally suggested that PA should be applied to the polychoric correlation matrix, especially when the variables are not symmetrical (Garrido et al., 2013; Timmerman et al., 2011). Although non-Gramian or nonpositive definite matrices can occur in the estimation of polychoric correlation matrices, smoothing the non-Gramian matrices with a procedure that always produces a Gramian matrix may resolve this problem (Garrido et al., 2013). Most studies on PA variants examined either continuous data or categorical data but not both together. A comprehensive evaluation of the PA procedures with such factor structures and data sets would have both theoretical and practical merits.

Simulation Study

The present study aims to evaluate the performance of PA procedures in determining the number of factors in a realistic or challenging data set where small eigenvalues are commonly present. We investigated the effect of small signal and noise eigenvalues on the accuracy of the PA procedures through a Monte Carlo simulation. Small signal eigenvalues were induced by manipulating (a) factor loading and (b) factor correlation. Small noise eigenvalues were induced by (a) adding minor error structure to a population correlation matrix and (b) imposing skewness or nonsymmetry that can produce difficulty factors. To accommodate a wide range of population structures in our study, we adopted a partially randomized parameter design to construct the population correlation matrices. We also examined the accuracy of the procedures in three different types of measurement scale.

Method

Design and Data Generation

The sample data set was generated in a different way depending on the measurement scale of the variable. The measurement scale was continuous, four-category ordinal, or binary. For each measurement scale, six design factors were manipulated as follows.

- **Number of factors (K).** The number of factors was one, two, four, or six.
- **Factor correlation (R).** For the multiple-factor model, that is, $K \geq 2$, factor correlation was manipulated as either *weak* or *strong*. The weak factor correlation was defined as a value within .0 and .3 while the strong factor correlation was defined as a value within .3 and .6.
- **Factor loading (L).** The factor loading was manipulated as either *low* or *high*. The low factor loading was defined as a value within .3 and .5 while the high factor loading was defined as a value within .5 and .7.
- **Population error (E).** Population error was either *present* or *not present*, as defined later.
- **Nonsymmetry (S).** Nonsymmetry or skewness was zero, one, or two.
- **Sample size (N).** The sample size was 100, 300, 500, 700, or 900.

The number of conditions for the one-factor model was 2 (L) $\times$ 2 (E) $\times$ 3 (S) $\times$ 5 (N) = 60 and the number of conditions for the multiple-factor model was 3(K) $\times$ 2(R) $\times$ 2(L) $\times$ 2(E) $\times$ 3(S) $\times$ 5 (N) = 360. As a result, a total of 420 conditions were generated for each measurement scale. The levels for the design factors were chosen so that they were representative of the range of values that are encountered in applied settings.

The first step of generating a sample data set began with the sampling of three types of random elements for the population model with a specified number of factors. The sampling of the random elements was done for each data set, so that the data sets within the same condition reflect less structured and more realistic population structures with some variations. First, the number of variables measuring a factor was sampled from the set {4, 6, 8} with equal probability so that the factors in the model can have different numbers of variables. Second, the loading values for each

measurement path were independently sampled from $U(.3, .5)$ for the low condition and $U(.5, .7)$ for the high condition and allowed to vary within a factor. Third, in case of the multiple-factor model, the correlation values for each pair of factors were independently sampled from $U(.0, .3)$ for the weak condition and $U(.3, .6)$ for the strong condition.

Once a population factor structure was set as above, the population correlation matrix $\mathbf{\Sigma}$ was constructed as follows (see Timmerman & Lorenzo-Seva, 2011). First, a reduced population correlation matrix $\mathbf{\Sigma}^*$ was computed as:

$$\mathbf{\Sigma}^* = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}} + (0.1)\mathbf{\Sigma}_{\mathbf{e}}^*, \quad (2)$$

where $\mathbf{\Lambda}$ is a $p \times k$ factor loading matrix, $\mathbf{\Lambda}^{\mathbf{T}}$ is the transpose of $\mathbf{\Lambda}$, $\mathbf{\Phi}$ is a $k \times k$ factor correlation matrix, and $\mathbf{\Sigma}_{\mathbf{e}}^*$ is a noise structure matrix. For models without population error, $\mathbf{\Sigma}_{\mathbf{e}}^*$ was a null matrix with 0 in every cell. For models with population error, the noise matrix was obtained by

$$\mathbf{\Sigma}_{\mathbf{e}}^* = \mathbf{\Lambda}_{\mathbf{e}}\mathbf{\Lambda}_{\mathbf{e}} / k,$$

where $\mathbf{\Lambda}_{\mathbf{e}}$ is a noise loading matrix of size $p \times 3k$ with each value sampled from $U(-1, 1)$. The denominator k is a normalization constant for the diagonal elements of $\mathbf{\Lambda}_{\mathbf{e}}\mathbf{\Lambda}_{\mathbf{e}}^{\mathbf{T}}$, the expected values of which are k. After computing the reduced population correlation matrix $\mathbf{\Sigma}^*$, the diagonal entries of $\mathbf{\Sigma}^*$ were replaced with unities to obtain $\mathbf{\Sigma}$, so that the communalities and unique variances summed to 1.0 for each variable. None of the generated $\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathbf{T}}$ had a diagonal value equal to or larger than 0.9.

For each of the 420 conditions, 100 sample data sets were simulated by drawing from a multivariate normal distribution with zero means and a specified population correlation $\mathbf{\Sigma}$. In the nonsymmetry condition, each data set was transformed to be skewed to a specified degree. For the continuous scale, the sample data was transformed by the Vale and Maurelli (1983) method. The level of nonsymmetry was manipulated by specifying a pair of skewness (γ_1) and excess kurtosis (γ_2) as (0, 0), (1, 1.5), or (2, 6). The values of each pair were set equal to the values of the chi-square distribution with $df = \infty$, 8 or 2, satisfying the mathematical bound $\gamma_2 \geq \gamma_1^2 - 2$ (Shohat, 1929). We used "mvrnonnorm" function from the "semTools" package in R for this purpose. For the categorical scales, each data set was first generated from a multivariate normal population with a specified $\mathbf{\Sigma}$ to serve as latent information. A set of prespecified thresholds was then applied to convert the data set into ordinal data set with a desired skewness level, that is, zero, one, or two. We used the thresholds provided in Garrido, Abad, and Ponsoda (2013). For all measurement scales, negative skewness values or sign-flipped thresholds were imposed on a half of the variables for each factor to simulate oppositely skewed variables, and thus induce difficulty factors.

Finally, a sample correlation matrix was obtained from each sample data. Pearson correlation was used for the data with continuous scale, and polychoric correlation was used for the data with categorical scale. Polychoric correlations were estimated with "polychoric" function from the "psych" package in R. In case of a nonpositive definite polychoric correlation matrix, smoothing was performed with the eigenvalue method described in Knol and Berger (1991), the default method in the function. A total of 13,345 (15.9%) matrices were smoothed out of 84,000 sample polychoric correlation matrices.

Analyses of Simulated Data

A total of 13 PA procedures listed in Table 1 were evaluated. For each sample data set, all procedures were performed as follows (see Table 2). First, PCA, PAF, and MRFA sample eigenvalues were obtained by performing respective communality estimation methods on the sample correlation matrix. For PCA eigenvalues, the diagonal elements were left unchanged. For PAF eigenvalues, the diagonal elements were replaced with SMC estimates. For MRFA eigenvalues, the communalities were estimated by MRFA. We used an archived Version 1.1.2 of the "DA.MRFA" package from CRAN repository to perform MRFA, suppling $p - 1$ as the rank to be used as in Timmerman and Lorenzo-Seva (2011). The three types of sample eigenvalues were used for the 12 procedures of "PA-" and "RPA-" families with matching types of communality estimators. PCA eigenvalues were used in CD. For the two PA-MRFA-ecv procedures, ECV values (eigenvalues divided by the sum of p eigenvalues) were used in place of eigenvalues for both the sample and random data sets.

For the eight procedures of the "PA-" family, only one set of 100 random data sets was generated for each sample data to construct the sampling distributions of eigenvalues or ECV values. This one set was then shared across the eight PA procedures to minimize unwanted random fluctuation from using different random data sets across the procedures. In case of the ordinal sample data set, the random data sets were also converted to ordinal scale as described in Table 2. The steps for converting a random data set into an ordinal scale and then again estimating the latent continuum may be regarded as redundant, but the steps were necessary to replicate the results of Green et al. (2016) in our preliminary simulations. From the 84,000 ordered categorical sample data sets, a total of 140 (0.17%) data sets did not converge when estimating a MRFA-reduced correlation matrix, for either the sample or a random data set. In these cases, a new sample data was generated by retrying from the first step of sampling the structural features of a population structure for all of the procedures.

For the four RPA procedures, the set of 100 random data sets used in the "PA-" family for each sample were used in the first $j = 0$ iteration. Also, only one set of 100 random data sets was generated and shared across the four RPA procedures in each iteration of increasing j. The j-factor EFA model was estimated with "factanal" function provided with R installation, using maximum likelihood (ML) estimation and performing iterative estimation on the communalities. An upper bound $j_{\max}$ was imposed because fitting a j-factor EFA model requires the model to be identified.[1] The CD procedure was performed with the sample data set of size $n \times p$. We used the code for R language obtained from the author's website. Default values were used for all arguments except the maximum number of tested factors, where $j_{\max}$ was imposed. The default value of 500 random data sets were generated in each increment of j.

[1] Separate bounds were imposed for the RPA-PCA and RPA-PAF to match their requirements as specified in Green et al. (2012). For a $p \times p$ correlation matrix, the number of available information is $a = p \times (p + 1)/2$ and the number of parameters in a j-factor EFA model is $b = (p \times j) + (j \times (j + 1)/2) + p - j^2$ (Brown, 2015). Solving the inequality $b \leq a$ for j yields the upper bound $j_{\max} = p - (0.5)((8p + 1)^{1/2}) + 0.5$ for RPA-PCA. The upper bound for RPA-PAF was specified as $j_{\max}^* = \min(j_{\max}, j_{\lambda \geq 0})$, where $j_{\lambda \geq 0}$ is the number of non-negative sample PAF eigenvalues.

The performance of the 13 procedures was evaluated and compared using three criteria. The percent correct (PC), the mean error (ME), and the root mean squared error (RMSE) in estimating the number of factors were calculated as in the following equations:

$$\mathrm{PC} = \frac{\sum C}{N_s} \times 100, \quad \text{for} \quad \mathrm{C} = \begin{Bmatrix} 1 \text{ if } \hat{\theta} = \theta \\ 0 \text{ if } \hat{\theta} \neq \theta \end{Bmatrix},$$

$$\mathrm{ME} = \frac{\sum (\hat{\theta} - \theta)}{N_s},$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum (\hat{\theta} - \theta)^2}{N_s}},$$

where N_s is the number of sample data sets used in the calculations, $\hat{\theta}$ is the estimated number of factors, and θ is the population number of factors (Garrido et al., 2013). The whole simulation was carried on R 3.2.2 software (R Core Team, 2015).

Results

For the sake of clear interpretation, we present only the results of seven out of the 13 procedures. Mostly, each procedure was chosen by comparing the overall performance of each pair of the methods which differ only in using the average or 95th percentile as a classification threshold. For example, PA-PAF-95 was selected because it performed better than PA-PAF-m. For the traditional PA procedures, both PA-PCA-m and PA-PCA-95 were included because they performed best among all procedures in most conditions. Among the four PA-MRFA procedures, only PA-MRFA-m was chosen because it consistently performed best. Table 3 shows the overall performance of the seven selected procedures.

All seven procedures were more accurate in the one-factor model than in the multiple-factor model, for example, PC(PA-PCA-m)$_{k=1}$ = 89.3 > PC(PA-PCA-m)$_{k\geq 2}$ = 65.1 for the continuous scale. In the one-factor model, the four PA families (i.e., PA-PCA-m, PA-PCA-95, PA-PAF-95, and PA-MRFA-m) performed better when the scale was continuous than when the scale was four-category or binary, for example, PC(PA-PCA-95)$_{k=1}$ = 94.2 versus 87.0 and 89.8, respectively. In contrast, the three procedures which update the population structure performed better when the scale was binary rather than continuous, for example, PC(RPA-PCA-95)$_{k=1}$ = 84.2 and 79.4, respectively. Regardless of the type of measurement scale, PA-PCA-95 was the most accurate procedure in the one-factor model, for example, PC(PA-PCA-95)$_{k=1}$ = 94.2 for the continuous scale. PA-MRFA-m and PA-PCA-m also showed relatively good performances in the one-factor model, for example, PC(PA-MRFA-m)$_{k=1}$ = 82.8 and PC(PA-PCA-m)$_{k=1}$ = 81.9 for the four-category scale. In the multiple-factor model, all procedures performed best when the measurement scale was continuous, next best when the scale was four-category, and worst when the scale was binary, for example, PC(CD)$_{k\geq 2}$ = 51.5, 44.4, and 36.5, respectively. The two traditional PA procedures (i.e., PA-PCA-m and PA-PCA-95) were the most accurate procedures in the multiple-factor model, for example, PC(PA-PCA-95)$_{k\geq 2}$ = 63.2, 51.4, and 45.5 when the scale was continuous, four-category, and binary, respectively. Overall, the use of communality estimates in PA, whether based on SMC or MRFA, did not increase the accuracy in estimating the number of factors, for example, PC(PA-PAF-95)$_{k=1}$ = 74.2 < PC(PA-PCA-95)$_{k=1}$ = 87.0 for the four-category scale. The three updating procedures (i.e., RPA-PCA-95, RPA-PAF-95, and CD) also failed to outperform the traditional PA procedures in most conditions, for example, PC(RPA-PCA-95)$_{k\geq 2}$ = 42.9, PC(RPA-PAF-95)$_{k\geq 2}$ = 38.8, and PC(CD)$_{k\geq 2}$ = 36.5 versus PC(PA-PCA-m)$_{k\geq 2}$ = 49.6 and PC(PA-PCA-95)$_{k\geq 2}$ = 45.5 for the binary scale.

To better understand the results, we examined the effects of the six design factors and their interactions on the proportion of correct estimates (PC/100) for the multiple-factor model. The analysis of variance (ANOVA) was conducted separately for each procedure and each measurement scale, resulting in a total of 21 ANOVAs. For each of the 360 conditions, the proportion of correct estimates was first calculated and then transformed by arcsine-square-root to deal with inherent heteroscedasticity in the bounded variables. Only up to three-way interactions were included in the analyses. To evaluate the effect size of the design factors in terms of the proportion of the variance accounted for by each factor, we calculated semipartial eta-squared (η_p^2), where $\eta_p^2 \geq .01$ is interpreted as a small effect, $\eta_p^2 \geq .06$ as a medium effect, and $\eta_p^2 \geq .14$ as a large effect (see Garrido et al., 2013). Table 4 presents the effect size of the six main effects and 3 two-way interactions that showed at least a medium effect in any procedure. Table 5, Table 6, and Table 7 present the PC, ME, and RMSE of the seven selected procedures across the levels of the six factors for the continuous scale, the four-category ordinal scale, and the binary scale, respectively. A detailed description of the main effects and notable interaction effects follows.

Population error had a substantial effect on the accuracy of the PA procedures, that is, $\bar{\eta}_p^2$ = .26, .08, and .02 when the scale was continuous, four-category, and binary, respectively. The effect size was most noticeable when the scale was continuous. All selected procedures were less accurate when population error was introduced (see Tables 5–7). For the continuous variable, the procedures most influenced were PA-PAF-95, RPA-PCA-95, and RPA-PAF-95, that is, η_p^2 = .52, .51, and .55, respectively. When population error was *absent*, the three procedures showed the highest accuracy, that is, PC(PA-PAF-95) = 78.1, PC(RPA-PCA-95) = 76.7, and PC(RPA-PAF-95) = 74.8. But in the *present* condition, their accuracy enormously dropped down to the lowest, that is, PC(PA-PAF-95) = 22.8, PC(RPA-PCA-95) = 20.7, and PC(RPA-PAF-95) = 15.9. A similar pattern of the effects was also found in the categorical scales. For example, PA-PAF-95, RPA-PCA-95, and RPA-PAF-95 were the most influenced by the population error when the scale was four-category, that is, η_p^2 = .25, .11, and .15, respectively. The accuracy of CD was also considerably affected by the population error, especially when the scale was continuous, that is, η_p^2 = .18. On the contrary, the traditional PA-PCA procedures were less affected by the population error than the alternative procedures except for PA-MRFA-m. The accuracy of PA-MRFA-m was similar between conditions with and without the population error for all measurement scales.

Population error seems to have increased the estimated number of factors. This can be inferred from the difference in ME between the *absent* and *present* conditions (Tables 5–7). The noise eigenvalues induced by the population error may increase common variance in the sample correlation matrix and thus increase the number of factors in the PA procedures. It also seems that the induced noise eigenvalues had an inflating effect on the com-

Table 3
Overall Performance of the Selected PA Variants

	Percent correct						Mean error						RMSE					
	K = 1			K ≥ 2			K = 1			K ≥ 2			K = 1			K ≥ 2		
Procedure	Cont.	4-C	2-C	Cont.	4-C	2-C	Cont.	4-C	2-C	Cont.	4-C	2-C	Cont.	4-C	2-C	Cont.	4-C	2-C
PA-PCA-m	89.3	81.9	84.0	65.1	54.1	49.6	.11	.13	.16	−.13	−.38	−.35	.37	.49	.49	.96	1.52	1.46
PA-PCA-95	94.2	87.0	89.8	63.2	51.4	45.5	.05	−.02	−.03	−.41	−.88	−1.05	.25	.36	.32	1.13	1.77	1.84
PA-PAF-95	78.8	74.2	76.1	50.4	38.1	30.1	.24	.14	.15	.56	.36	1.79	.58	.59	.72	1.72	3.61	7.39
PA-MRFA-m	92.6	82.8	83.2	52.9	46.8	44.8	.02	.00	.04	−.45	−.61	−.63	.29	.44	.45	1.10	1.57	1.50
RPA-PCA-95	79.4	78.6	84.2	48.7	44.2	42.9	.24	.08	.04	.53	−.62	−.95	.58	.52	.42	1.87	1.98	1.93
RPA-PAF-95	73.1	73.9	80.4	45.4	40.2	38.8	.35	.15	.06	.75	−.56	−1.03	.74	.62	.50	2.14	2.13	2.01
CD	75.6	83.6	89.0	51.5	44.4	36.5	.37	.23	.14	−.18	−1.09	−1.53	.82	.61	.47	1.59	2.12	2.45

Note. RMSE = root mean squared error; K = number of factors; Cont. = continuous scale; 4-C = four-category scale; 2-C = binary scale. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; PCA = principal component analysis; PAF = principal axis factoring; MRFA = minimum rank factor analysis; m = mean; 95 = 95th percentile.

munality estimates in the PA-PAF, but not in the PA-MRFA. The noise eigenvalues also seem to have increased inaccuracy in estimating the factor structures updated in the RPA and CD. Such negative effects may cause more serious concern when the procedures are applied to Pearson correlations of continuous variables than to polychoric correlations of categorical variables.

Nonsymmetry had less-than-small effects for the continuous variable but small-to-medium effects for the categorical variables, that is, $\bar{\eta}_p^2$ = .08 and .02 when the scale was four-category and

Table 4
Effect Size of the Selected PA Variants in the Multiple-Factor Model by Measurement Scale

	Main and interaction effects								
Procedure	E	S	R	L	K	N	E × N	R × L	R × K
				Continuous					
PA-PCA-m	.04		*.14*	**.24**	**.19**	*.09*	.02	.02	*.07*
PA-PCA-95			**.20**	**.19**	**.24**	**.14**			*.06*
PA-PAF-95	**.52**		.02	*.11*	*.09*		**.15**		
PA-MRFA-m			**.37**	*.12*	**.27**	.02		*.11*	.04
RPA-PCA-95	**.51**		.01	.05	*.09*	.01	**.15**		
RPA-PAF-95	**.55**			.02	.05		**.15**		
CD	**.18**		.04	**.16**	*.10*	*.09*	*.08*		.01
Average	**.26**		*.11*	*.12*	**.15**	.05	*.08*	.02	.03
				Four-Category					
PA-PCA-m	.01	.05	*.10*	**.19**	**.18**	**.20**		.02	.04
PA-PCA-95		.05	**.17**	*.14*	**.21**	**.26**			.04
PA-PAF-95	**.25**	*.06*		*.08*	**.15**	**.17**	*.11*		
PA-MRFA-m		.04	**.25**	*.09*	**.22**	*.12*		*.11*	.02
RPA-PCA-95	*.11*	.02	.04	*.11*	*.14*	**.22**	*.07*		
RPA-PAF-95	**.15**	.01	.02	*.06*	*.11*	**.21**	*.09*		
CD	.02	.06	*.10*	**.21**	**.25**	**.16**	.01		
Average	*.08*	.04	*.10*	*.13*	**.18**	**.19**	.04	.02	.02
				Binary					
PA-PCA-m		.03	*.12*	**.24**	**.24**	**.15**		.05	.04
PA-PCA-95		.03	**.18**	**.17**	**.27**	**.21**		.02	.02
PA-PAF-95	*.11*	*.07*	.02	*.07*	**.33**	*.10*	*.07*		
PA-MRFA-m		.01	**.26**	*.12*	**.30**	*.10*		*.10*	.03
RPA-PCA-95	.02	.03	.05	**.24**	**.23**	**.21**	.02		
RPA-PAF-95	.02	.03	.04	**.22**	**.22**	**.19**	.02		
CD		*.06*	*.10*	**.22**	**.34**	**.15**			
Average	.02	.04	*.11*	**.18**	**.28**	**.16**	.02	.03	.01

Note. Tabled values are semipartial eta squared (η_p^2) obtained from the analysis of variance on the proportion of correct estimates transformed by arcsine-square-root. Values greater than small effects ($\eta_p^2 > .01$) are displayed. Large effects ($\eta_p^2 > .14$) are shown in boldface. Medium effects ($\eta_p^2 > .06$) are shown in italics. E = population error; S = nonsymmetry (skewness); R = factor correlation; L = factor loading; K = number of factors; N = sample size. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; PCA = principal component analysis; PAF = principal axis factoring; MRFA = minimum rank factor analysis; m = mean; 95 = 95th percentile.

Table 5
Performance of the Selected PA Variants in the Multiple-Factor Model for the Continuous Variable

Procedure	Population error		Nonsymmetry[a]			Factor correlation		Factor loading		Number of factors			Sample size				
	No	Yes	0	1	2	Weak	Strong	High	Low	2	4	6	100	300	500	700	900
								Percent correct									
PA-PCA-m	71.7	58.4	64.9	65.6	64.7	76.5	53.6	81.4	48.7	83.0	66.5	45.7	47.2	64.5	69.4	71.5	72.6
PA-PCA-95	66.1	60.3	63.4	63.6	62.6	78.0	48.5	77.8	48.7	85.1	63.0	41.5	38.6	62.8	68.9	72.2	73.5
PA-PAF-95	78.1	22.8	50.7	50.3	50.4	55.1	45.8	61.6	39.2	63.9	49.4	38.0	46.4	54.9	50.9	49.7	50.2
PA-MRFA-m	53.7	52.1	53.1	53.8	51.9	72.4	33.4	64.2	41.6	76.1	48.9	33.8	43.1	54.1	55.7	56.0	55.8
RPA-PCA-95	76.7	20.7	48.7	48.6	48.8	53.3	44.1	56.8	40.6	61.8	47.3	37.0	43.9	53.9	49.7	48.1	47.9
RPA-PAF-95	74.8	15.9	45.4	46.0	44.7	49.6	41.1	51.5	39.2	55.6	44.8	35.7	42.2	48.6	45.6	45.2	45.2
CD	67.3	35.6	47.0	50.5	56.8	58.2	44.8	65.5	37.4	64.6	51.9	37.8	33.7	54.7	56.0	56.7	56.2
								Mean error									
PA-PCA-m	−.37	.11	−.12	−.12	−.14	.18	−.44	−.23	−.02	.14	−.03	−.49	−.58	−.16	−.03	.04	.09
PA-PCA-95	−.58	−.25	−.41	−.40	−.43	−.06	−.77	−.31	−.52	.01	−.32	−.93	−1.23	−.46	−.23	−.11	−.04
PA-PAF-95	−.24	1.37	.58	.58	.54	.75	.38	.51	.62	.44	.65	.60	−.75	.42	.84	1.07	1.23
PA-MRFA-m	−.50	−.41	−.43	−.43	−.50	.03	−.94	−.53	−.38	−.06	−.42	−.87	−.68	−.39	−.39	−.41	−.40
RPA-PCA-95	−.32	1.39	.55	.56	.49	.72	.34	.91	.16	.48	.62	.50	−.93	.21	.77	1.15	1.46
RPA-PAF-95	−.31	1.81	.77	.78	.70	.92	.58	1.29	.21	.69	.87	.70	−.87	.38	1.02	1.45	1.78
CD	−.62	.26	−.02	−.05	−.46	.08	−.44	.23	−.58	.44	−.01	−.96	−1.48	−.40	.05	.36	.59
								Root mean squared error									
PA-PCA-m	.90	1.02	.96	.95	.97	.74	1.14	.63	1.20	.51	.82	1.36	1.44	.91	.78	.74	.74
PA-PCA-95	1.16	1.09	1.12	1.11	1.15	.73	1.41	.76	1.40	.41	.92	1.67	1.88	1.07	.82	.73	.68
PA-PAF-95	.88	2.26	1.75	1.72	1.68	1.67	1.76	1.27	2.07	.96	1.64	2.29	1.62	1.27	1.61	1.88	2.08
PA-MRFA-m	1.11	1.09	1.07	1.08	1.15	.71	1.39	.97	1.22	.55	.96	1.55	1.45	1.02	.98	.98	.99
RPA-PCA-95	.97	2.46	1.90	1.91	1.80	1.86	1.89	1.86	1.88	1.04	1.80	2.49	1.80	1.22	1.60	2.04	2.45
RPA-PAF-95	.98	2.86	2.19	2.19	2.03	2.13	2.14	2.35	1.90	1.33	2.13	2.72	1.78	1.39	1.92	2.43	2.86
CD	1.39	1.76	1.66	1.54	1.55	1.36	1.79	1.22	1.89	1.05	1.36	2.15	2.30	1.38	1.24	1.30	1.47

Note. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; PCA = principal component analysis; PAF = principal axis factoring; MRFA = minimum rank factor analysis; m = mean; 95 = 95th percentile.
[a] Skewness paired with a specified kurtosis.

binary, respectively. When the scale was categorical, the seven procedures became less accurate as the nonsymmetry increased (see Tables 6 and 7). The accuracy decreased more sharply, as the skewness increased gradually, for example, for the four-category scale, PC(PA-PAF-95) = 43.8, 43.5, and 26.9 when nonsymmetry was 0, 1, and 2, respectively. PA-PAF-95 and CD were the procedures most affected by the nonsymmetry, for example, η_p^2(PA-PAF-95) = .07 and η_p^2 (CD) = .06 for the binary scale. Nonsymmetry contributed to increase in positive bias for PA-PAF-95 and negative bias for the rest of the procedures. For example, as skewness increased from 0 to 2 in the binary scale, the ME of PA-PAF-95 increased from 1.02 to 2.60 but the ME of CD decreased from −1.29 to −1.86. The small noise eigenvalues induced by nonsymmetry seem to have more effect on categorical variables than continuous variables.

The factor correlation contributed a lot to lack of the accuracy for all procedures in all measurement scales, that is, $\bar{\eta}_p^2$ = .11, .10, and .11 for the continuous, four-category, and binary variables, respectively. For all scales, the seven procedures were less accurate in the *strong* condition than in the *weak* condition (Tables 5–7). PA-MRFA-m was found to have the largest effect size of the factor correlation among all procedures in all measurement scales, that is, η_p^2 = .37, .25, and .26 for the continuous, four-category, and binary scale, respectively. The accuracy of PA-PCA-95 was also largely affected by the factor correlation in all scales, that is, η_p^2 = .20, .17, and .18 for the continuous, four-category, and binary scale, respectively. The two procedures showed the largest difference in PC between the *weak* and *strong* conditions, for example, PC(PA-MRFA-m) = 72.4 versus 33.4 and PC(PA-PCA-95) = 78.0 versus 48.5 for the continuous variable. In all measurement scales, PA-PCA-m was the most accurate in the *strong* condition and PA-PCA-95 was the most accurate in the *weak* condition. The ME was negatively oriented in the *strong* condition than in the *weak* condition in all procedures. The strong factor correlation seems to have contributed to a reduction in the estimated number of factors because it possibly made the last few signal eigenvalues too small to be detected.

Factor loading also had a substantial effect on the accuracy of all procedures in all measurement scales. The average η_p^2 was .12 for the continuous scale, .13 for the four-category scale, and .18 for the binary scale, indicating that the effect of the factor loading was most evident when the scale was binary. In all scales, the accuracy of all procedures was greater in the *high* condition than in the *low* condition. The accuracy of the traditional PA procedures was largely affected by the factor loadings in all scales, for example, in PA-PCA-m, η_p^2 = .24, .19, and .24 for the continuous, four-category, and binary scale, respectively. CD also showed a large effect size in all scales, that is, η_p^2 = .16, .21, and .22 for the continuous, four-category, and binary scale, respectively. The traditional PA procedures were the most accurate in the *high* condition for all scales. Both procedures were also the best in the *low* condition with one exception, where PA-MRFA-m was the most

Table 6
Performance of the Selected PA Procedures in the Multiple-Factor Model for the Four-Category Ordinal Variable

Procedure	Population error		Nonsymmetry[a]			Factor correlation		Factor loading		Number of factors			Sample size				
	No	Yes	0	1	2	Weak	Strong	High	Low	2	4	6	100	300	500	700	900
								Percent correct									
PA-PCA-m	58.2	49.9	60.9	58.4	42.9	64.5	43.6	69.0	39.1	73.0	54.0	35.2	26.3	49.4	60.8	65.5	68.2
PA-PCA-95	52.1	50.8	58.2	55.2	40.9	66.7	36.2	64.9	37.9	74.4	49.3	30.6	17.5	47.7	59.1	64.4	68.4
PA-PAF-95	56.1	20.0	43.8	43.5	26.9	39.9	36.2	47.5	28.6	55.3	34.8	24.1	14.6	40.5	44.9	44.7	45.6
PA-MRFA-m	47.0	46.7	51.9	50.0	38.5	63.0	30.6	56.9	36.7	67.3	43.2	30.0	25.4	46.4	51.8	55.4	55.1
RPA-PCA-95	57.1	31.2	47.7	47.0	37.9	50.9	37.5	55.4	32.9	61.3	41.8	29.4	15.4	43.8	53.1	54.3	54.2
RPA-PAF-95	55.4	25.1	41.7	43.6	35.4	45.4	35.0	48.9	31.5	55.2	38.0	27.4	12.2	42.0	48.4	49.1	49.4
CD	49.2	39.5	52.6	47.3	33.2	55.6	33.1	60.6	28.2	67.3	40.7	25.1	20.0	42.3	49.0	53.9	56.4
								Mean error									
PA-PCA-m	−.59	−.18	−.21	−.25	−.69	−.01	−.76	−.45	−.32	.08	−.28	−.94	−1.62	−.10	−.13	−.06	−.01
PA-PCA-95	−1.02	−.75	−.61	−.70	−1.34	−.48	−1.29	−.66	−1.10	−.20	−.81	−1.63	−2.51	−.78	−.53	−.35	−.23
PA-PAF-95	−.40	1.13	.10	.40	.59	.67	.06	.64	.09	.20	.24	.65	−2.65	1.04	1.05	1.12	1.26
PA-MRFA-m	−.67	−.56	−.35	−.51	−.98	−.17	−1.06	−.74	−.49	−.12	−.54	−1.18	−1.53	−.42	−.40	−.35	−.37
RPA-PCA-95	−.98	−.26	−.12	−.30	−1.43	−.37	−.86	−.18	−1.05	−.01	−.54	−1.31	−2.66	−.94	−.18	.20	.49
RPA-PAF-95	−1.03	−.09	−.12	−.22	−1.34	−.36	−.76	−.05	−1.06	.12	−.49	−1.30	−2.96	−.82	−.06	.37	.68
CD	−1.26	−.91	−.62	−1.07	−1.58	−.67	−1.51	−.55	−1.63	.05	−.98	−2.34	−2.21	−1.28	−.93	−.61	−.41
								Root mean squared error									
PA-PCA-m	1.50	1.54	1.09	1.19	2.08	1.42	1.61	1.22	1.77	.73	1.35	2.14	2.71	1.29	1.02	.89	.84
PA-PCA-95	1.81	1.72	1.36	1.47	2.32	1.49	2.00	1.45	2.04	.71	1.55	2.54	3.18	1.56	1.18	.98	.86
PA-PAF-95	2.84	4.25	2.15	2.39	5.37	3.89	3.31	4.11	3.03	1.09	2.57	5.60	3.79	5.24	3.47	2.34	2.45
PA-MRFA-m	1.58	1.55	1.14	1.26	2.11	1.39	1.73	1.41	1.70	.80	1.39	2.19	2.72	1.29	1.10	1.00	1.00
RPA-PCA-95	1.86	2.10	1.68	1.68	2.49	1.84	2.12	1.70	2.23	.92	1.77	2.80	3.29	1.92	1.28	1.24	1.42
RPA-PAF-95	1.94	2.29	2.09	1.77	2.46	2.03	2.22	1.94	2.30	1.08	1.95	2.93	3.54	1.87	1.36	1.43	1.65
CD	2.13	2.10	1.65	2.10	2.51	1.64	2.50	1.61	2.52	.75	1.69	3.16	2.86	2.19	1.95	1.72	1.62

Note. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; PCA = principal component analysis; PAF = principal axis factoring; MRFA = minimum rank factor analysis; m = mean; 95 = 95th percentile.
[a] Skewness defined from specified thresholds.

accurate for the binary scale. RPA-PCA-95, RPA-PAF-95, and CD showed a substantial negative bias in the *low* condition for the categorical scales.

Overall, the number of factors was found to be the most influential design factor, that is, $\bar{\eta}_p^2$ = .15, .18, and .28 for the continuous, four-category, and binary variable, respectively. The increase in the number of factors remarkably reduced the accuracy of the seven procedures in all scales. The accuracy of the traditional PA procedures was largely influenced by the number of factors in all measurement scales, for example, η_p^2 (PA-PCA-95) = .24, .21, and .27 for the continuous, four-category, and binary scale, respectively. PA-MRFA-m also had a large effect size of the number of factors on all scales, that is, η_p^2 = .27, .22, and .30 for the continuous, four-category, and binary scale, respectively. PA-PCA-m was the most accurate regardless of the number of factors for all scales. For the continuous scale, the negative bias increased as the number of factors increased in PA-PCA-m, PA-PCA-95, PA-MRFA-m, and CD. For the categorical scales, as the number of factors increased, the negative bias increased in all procedures except for PA-PAF-95, for which the increase in the number of factors led to increase of the positive bias on all scales.

Sample size had a substantial effect on the accuracy of all procedures in all measurement scales. The average η_p^2 was .05 for the continuous variable, .19 for the four-category variable, and .16 for the binary variable. The effect of the sample size was more evident when the scale was categorical and the accuracy in conditions with the sample size of 100 was markedly low, for example, PC(PA-PCA-95) = 16.1 and PC(PA-PAF-95) = 14.8 for the binary scale. The accuracy of the traditional PA procedures was largely influenced by the sample size in all measurement scales, for example, η_p^2 (PA-PCA-95) = .14, .26, and .21 for the continuous, four-category, and binary scale, respectively. As the sample size increased, the accuracy of both procedures consistently increased in all measurement scales. PA-PCA-m was the most accurate in all sample size conditions with a few exceptions where PA-PCA-95 performed slightly better, for example, PC(PA-PCA-m) = 72.6 and PC(PA-PCA-95) = 73.5 when the sample size is 900 for the continuous scale. For the categorical scales, PA-MRFA-m had a smaller effect than the other procedures, that is, η_p^2 = .12 and .10 for the four-category and binary scale, respectively. As the sample size increased, the accuracy of PA-MRFA-m did not increase as much as the accuracy of the other procedures. PA-PAF-95, RPA-PCA-95, RPA-PAF-95, and CD had a medium to large effect of sample size for the categorical variables, for example, for the four-category scale, η_p^2 = .17, .22, .21, and .16, respectively. The accuracy of the four procedures generally increased as the sample size increased for the categorical variables.

For the continuous scale, however, the accuracy of the alternative PA procedures did not behave in a typical way as the sample size increased. For example, the PC of RPA-PCA-95 was highest

Table 7
Performance of the Selected PA Variants in the Multiple-Factor Model for the Binary Variable

Procedure	Population error		Nonsymmetry[a]			Factor correlation		Factor loading		Number of factors			Sample size				
	No	Yes	0	1	2	Weak	Strong	High	Low	2	4	6	100	300	500	700	900
								Percent correct									
PA-PCA-m	52.3	46.9	55.3	51.2	42.5	60.7	38.6	66.7	32.5	71.2	48.1	29.6	26.9	45.7	54.3	59.2	62.0
PA-PCA-95	45.6	45.3	51.9	47.6	36.9	61.0	29.9	60.7	30.2	71.1	41.2	24.1	16.1	40.0	51.7	57.4	61.9
PA-PAF-95	40.4	19.9	38.1	31.3	21.0	34.1	26.2	38.1	22.2	51.7	24.7	14.0	14.8	30.7	34.9	35.1	35.2
PA-MRFA-m	45.2	44.4	48.7	46.1	39.7	60.5	29.1	56.7	32.9	67.7	41.1	25.6	26.6	42.0	49.5	51.8	54.1
RPA-PCA-95	48.8	37.0	47.6	45.5	35.5	50.4	35.3	59.7	26.1	63.8	39.4	25.6	15.5	40.3	50.4	53.6	54.7
RPA-PAF-95	44.7	33.0	43.2	41.7	31.6	45.8	31.9	54.0	23.7	59.0	35.0	22.6	13.8	35.0	46.2	49.7	49.5
CD	38.3	34.7	45.3	39.6	24.6	47.2	25.8	52.7	20.3	63.0	30.4	16.1	14.0	32.3	40.9	45.6	49.7
								Mean error									
PA-PCA-m	−.53	−.17	−.27	−.33	−.45	.13	−.83	−.46	−.24	.15	−.24	−.96	−1.11	−.39	−.16	−.07	−.02
PA-PCA-95	−1.16	−.94	−.81	−.96	−1.38	−.60	−1.50	−.72	−1.38	−.23	−.96	−1.96	−2.32	−1.20	−.77	−.55	−.42
PA-PAF-95	.98	2.61	1.02	1.76	2.60	2.13	1.46	3.26	.33	.48	1.95	2.95	−2.21	1.54	2.34	3.25	4.05
PA-MRFA-m	−.70	−.56	−.52	−.60	−.78	−.11	−1.15	−.68	−.58	−.04	−.53	−1.32	−1.19	−.65	−.48	−.44	−.39
RPA-PCA-95	−1.14	−.76	−.54	−.85	−1.47	−.69	−1.22	−.32	−1.59	−.12	−.86	−1.88	−2.40	−1.21	−.69	−.36	−.12
RPA-PAF-95	−1.25	−.81	−.54	−.93	−1.62	−.78	−1.28	−.36	−1.70	−.13	−.94	−2.02	−2.43	−1.35	−.78	−.42	−.16
CD	−1.63	−1.44	−1.29	−1.45	−1.86	−1.06	−2.01	−.92	−2.15	−.14	−1.40	−3.06	−2.54	−1.77	−1.35	−1.10	−.91
								Root mean squared error									
PA-PCA-m	1.45	1.47	1.25	1.39	1.71	1.26	1.64	1.06	1.78	.75	1.27	2.06	2.14	1.53	1.25	1.08	1.02
PA-PCA-95	1.89	1.79	1.57	1.74	2.16	1.49	2.14	1.35	2.23	.66	1.58	2.70	2.91	1.91	1.47	1.24	1.09
PA-PAF-95	6.69	8.02	5.07	7.02	9.42	7.41	7.37	9.22	4.90	2.13	6.66	10.71	3.07	7.75	7.46	7.94	9.20
PA-MRFA-m	1.52	1.48	1.30	1.43	1.74	1.15	1.78	1.24	1.72	.71	1.28	2.15	2.13	1.55	1.30	1.19	1.09
RPA-PCA-95	1.94	1.92	1.64	1.81	2.28	1.74	2.10	1.30	2.40	.75	1.65	2.80	3.00	1.99	1.55	1.32	1.22
RPA-PAF-95	2.02	2.00	1.71	1.87	2.38	1.84	2.17	1.38	2.49	.83	1.73	2.90	3.01	2.09	1.63	1.46	1.39
CD	2.47	2.44	2.27	2.39	2.67	1.98	2.85	1.93	2.88	.68	1.99	3.69	3.12	2.59	2.29	2.11	1.97

Note. PA = parallel analysis; RPA = revised parallel analysis; CD = comparison data method; PCA = principal component analysis; PAF = principal axis factoring; MRFA = minimum rank factor analysis; m = mean; 95 = 95th percentile.
[a] Skewness defined from specified thresholds.

when the sample size was 300 and thereafter decreased as the sample size increased, that is, PC(RPA-PCA-95) = 43.9, 53.9, 49.7, 48.1, and 47.9 for N = 100, 300, 500, 700, and 900, respectively. In fact, a medium to large effect of the interaction of sample size by population error was found for PA-PAF-95, RPA-PCA-95, RPA-PAF-95, and CD, that is, η_p^2 = .15, .15, .15, and .08, respectively. Figure 3 demonstrates the differences in PC of the selected PA procedures across the different levels of sample size in each combination of the factor correlation and factor loading conditions when the population error was (a) *absent* or (b) *present* for the continuous scale. In Figure 3, PA-PCA-95 and RPA-PAF-95 are not displayed for brevity because each was less accurate than and had a similar pattern to PA-PCA-m and RPA-PCA-95, respectively. Without population error, all displayed procedures except for PA-MRFA-m showed increased accuracy as the sample size increased. When the population error was *present*, however, the alternative PA procedures showed unexpected patterns of change in accuracy. For example, the accuracy of PA-PAF-95 and RPA-PCA-95 decreased as the sample size increased when the factor loading was high. Both procedures behaved strangely also when the factor loading was low. CD also exhibited nonsystematic changes with increasing sample size in the *present* condition. In contrast, PA-PCA-m showed a systematic increase in accuracy as the sample size increased when the factor loading was high in the *present* condition. Similar patterns of the interactions were also found in the four-category and binary scale but the effect sizes were much smaller than in the continuous scale, for example, η_p^2 (RPA-PCA-m) = .07, and .02 for the four-category and binary scale, respectively. Figure 4 demonstrates the differences in PC of the five selected PA procedures across the different levels of sample size in each combination of the factor correlation and factor loading conditions for the (a) four-category scale and (b) the binary scale.

A notable interaction effect of factor correlation by factor loading was found for PA-MRFA-m, that is, η_p^2 = .11, .11, and .10 for the continuous, four-category, and binary scale, respectively. PA-MRFA-m performed well when both the factor correlation was weak and the factor loading was high in all scales. However, the accuracy of the procedure dropped a lot when either the factor correlation was strong or the factor loading was low (see Figure 3 and Figure 4). The interaction indicates that just one source of small signal eigenvalue, whether it is a strong factor correlation or a low factor loading, is enough to decrease the accuracy of PA-MRFA. A similar pattern of interaction was also found in PA-PCA-m, but the effect size of the procedure was much smaller than PA-MRFA-m, that is, η_p^2 = .02, .02, and .05 for the continuous, four-category, and binary scale, respectively. PA-PCA-m and PA-MRFA-m also had a small to medium interaction effect of the number of factors by factor correlation, for example, η_p^2 (PA-PCA-m) = .07, .04, and .04 for the continuous, four-category, and binary scale, respec-

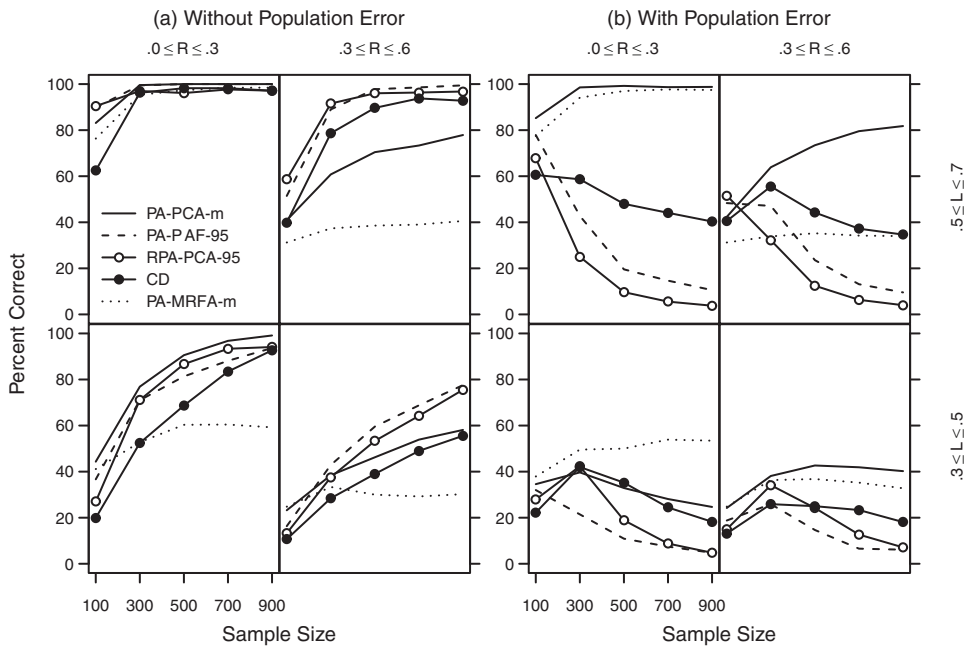

Figure 3. Percent correct of the selected PA procedures in conditions of Sample Size × Factor Correlation × Factor Loading for the continuous variable. Results without population error (a) and with population error (b) are presented separately. R = factor correlation; L = factor loading.

tively. The interaction indicates that the negative effect of the number of factors on the accuracy of the two procedures is greater when the factor correlation is stronger than when it is weaker.

Conclusion

Previous studies have pointed out that parallel analysis lacks theoretical justification because it does not use proper communal-

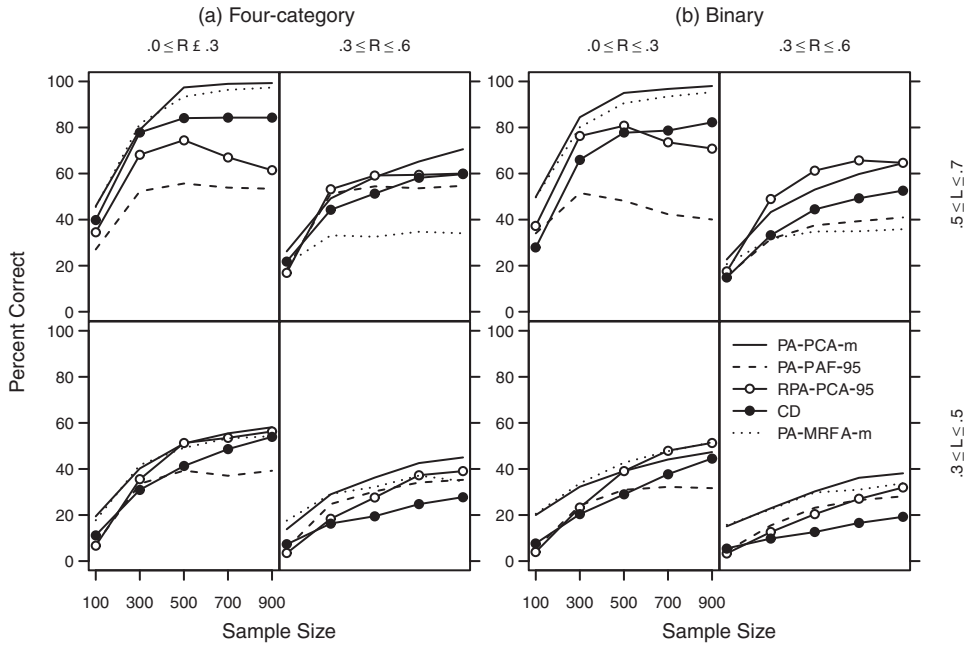

Figure 4. Percent correct of the selected PA procedures in conditions of Sample Size × Factor Correlation × Factor Loading for the categorical variable. Results for the four-category variable (a) and for the binary variable (b) are presented separately. R = factor correlation; L = factor loading.

ity estimates nor consider the constrained relation in eigenvalues. Alternative PA procedures to address the problems have been proposed in the literature, and the results in the respective studies suggest that the alternatives provide a benefit over the traditional parallel analysis in certain conditions. However, in order to reach a general conclusion about the performance of PA procedures, they must be investigated under a wide range of factor structures with small signal and noise eigenvalues. The current study examined performance of 13 PA variants across various forms of factor structure, focusing on the capability of the procedures in correctly detecting small factor-representing eigenvalues and correctly rejecting small noise eigenvalues. The procedures were investigated in both continuous data sets and ordered categorical data sets, using appropriate types of Pearson or polychoric correlations.

Our results show that Horn's original PA generally outperforms the alternative PA procedures, especially when the population factor structure has model error or trivial factors. PA-PCA-95 was the most accurate for the one-factor model, and PA-PCA-m was the most accurate for the multiple-factor model, regardless of whether the measurement scale was continuous or categorical. However, the traditional PA procedures were still not sufficiently accurate in the conditions with small signal eigenvalues, which was induced by strong factor correlation and low factor loadings in our design. The lack of a sufficient accuracy was especially more prominent as the number of factors increased. The small signal eigenvalue may contribute to decreasing the accuracy of parallel analysis as the number of factors increases. In fact, regardless of the number of factors, the accuracy of PA-PCA-m for the multiple factor model was satisfactory in condition with high factor loading and weak factor correlation, where the small signal eigenvalue is minimally induced (see Figures 3–4).

The alternative PA procedures outperformed or had comparable accuracy with the original PA in only a few conditions. Parallel analysis using SMCs as communality estimates outperformed the original PA in condition with strong factor correlation (Crawford et al., 2010; Green et al., 2012) but this was no longer true once model error was involved. Similarly, revised parallel analysis outperformed Horn's PA in condition with strong factor correlation (Green et al., 2012, 2016) but it failed to detect a small eigenvalue induced by model error as a noise. CD showed better performance than the original PA when both factor correlation was strong and factor loading was high but this was true only if population error was not involved. Parallel analysis using MRFA communality estimates showed a comparable accuracy with the original PA in condition with weak factor correlation and high factor loading (Timmerman & Lorenzo-Seva, 2011) but it performed badly when factors are highly correlated (Garrido et al., 2013).

In summary, our results show that the proposed PA alternatives generally do not succeed to address theoretical limitations of Horn's parallel analysis. The results of the current study suggest that the original PA procedure should be the method of choice to determine the number of factors using eigenvalues. Unless the factor model under study is expected to fit exactly in the population, the variables used in PA must be reliable and valid, enough in number, and have a single meaning to avoid a small signal eigenvalue. In case of the categorical variable, PA must be performed on a data set with a larger sample size and for the factor model with a smaller number of factors, because PA is affected a lot by the number of factors and sample size.

In a sample correlation matrix, a similar set of eigenvalues can be obtained from different population factor structures. For example, eigenvalues from a correlated-factor structure could yield a similar pattern to those from an orthogonal-factor structure with the same number of factors. The two factor structures in Figure 5a have different factor correlations and factor loadings, but they produce nearly identical sample eigenvalue distributions because the two structures are essentially congruent after rotation. Furthermore, when factors are strongly correlated the representing signal eigenvalues become so small that they become indistinguishable with noise eigenvalues, and even become reproducible with a smaller number of factors and weak factor correlations. In Figure 5b, the fourth sample eigenvalues of the four-correlated-factor structure are indistinguishable with the fourth sample eigenvalues of the three-orthogonal-factor structure.

Concerning this lack of decidability and the heuristic nature of eigenvalue-based decision procedures in dimensionality assessment, the estimate obtained from PA or any eigenvalue-based procedure should not be interpreted as a fixed estimate. With the lack of decidability in mind, researchers may first obtain an estimate of the number of factors k from PA-PCA, and then considering $(k + 1)$, k, and $(k - 1)$ as viable candidates for the optimal number of factors. This is because the given set of sample eigenvalues may be produced from a k-factor structure as well as from a $(k + 1)$-factor or a $(k - 1)$-factor structure, as demonstrated in Figure 5. As such, it is important that researchers should consult with the interpretability of the factor structures and compare viable models with different number of factors in terms of their interpretational validity, typically after rotation (Fabrigar et al., 1999; Floyd & Widaman, 1995; Reise, Waller, & Comrey, 2000).

The scope of the current study is limited in many ways. While our study design covered a wide range of realistic conditions, these conditions are not comprehensively representative of the vast range of the factor structures in real data sets. For example, our conditions produce data sets with limited degrees of nonsymmetry. The current study did not investigate more analytic eigenvalue-based approaches based on random matrix theory recently proposed by Saccenti and Timmerman (2017) and Braeken and van Assen (2017). Other approaches of dimensionality assessment not based on the eigenvalues of correlation matrix, for example, likelihood ratio test, were not examined in the current study, either. Comparisons with such approaches under comprehensive and realistic conditions may enhance our understanding of the nature of dimensionality assessment.

The conditions examined here only represents some portion of possible configuration of common factor models. Therefore, readers are advised to be cautious in generalizing the current results into other models, such as factor models with correlated measurement errors or cross loadings, or bifactor models. Applying parallel analysis procedures onto data sets where such modeling is more appropriate may yield an inaccurate estimate. As such, while having a well-functioning estimation procedure is certainly of merit in psychological research, it would nonetheless still be important to investigate that for a given data set, whether the model structure in question is theoretically appropriate and conceptually interpretable.

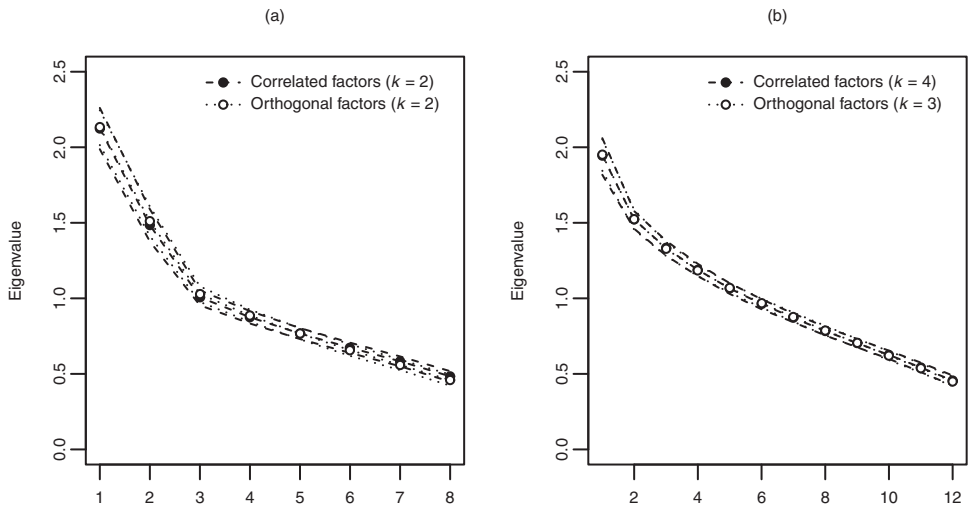

Figure 5. Eigenvalue similarity between different factor structures. Panel (a) illustrates the structures with two correlated factors [J = 4, L = .5, R = .3] and two orthogonal factors [J = 4, L = (.588, .409)]. Panel (b) illustrates the structures with four correlated factors [J = 3, L = .4, R = .3] and three orthogonal factors [J = 4, L = (.496, .299, .250)]. J = the number of variables for each factor; L = factor loadings for each factor (no cross-loadings were specified); R = factor correlation for each pair of the factors. The eigenvalue distributions were constructed from 1,000 datasets with 100 observations each. The three lines respectively represent the 25th percentile, the mean, and the 75th percentile of the distributions.

References

Achim, A. (2017). Testing the number of required dimensions in exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 13,* 64–74. http://dx.doi.org/10.20982/tqmp.13.1.p064

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

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27,* 509–540. http://dx.doi.org/10.1207/s15327906mbr2704_2

Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70,* 885–901. http://dx.doi.org/10.1177/0013164410379332

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

Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research, 27,* 387–415. http://dx.doi.org/10.1207/s15327906mbr2703_5

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286–299. http://dx.doi.org/10.1037/1040-3590.7.3.286

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. *Psychological Methods, 18,* 454–474. http://dx.doi.org/10.1037/a0030005

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55,* 377–393. http://dx.doi.org/10.1177/0013164495055003002

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

Green, S. B., Levy, R., Thompson, M. S., Lu, M., & Lo, W. J. (2012). A proposed solution to the problem with using completely random data to assess the number of factors with parallel analysis. *Educational and Psychological Measurement, 72,* 357–374. http://dx.doi.org/10.1177/0013164411422252

Green, S. B., Redell, N., Thompson, M. S., & Levy, R. (2016). Accuracy of revised and traditional parallel analyses for assessing dimensionality with binary data. *Educational and Psychological Measurement, 76,* 5–21. http://dx.doi.org/10.1177/0013164415581898

Green, S. B., Thompson, M. S., Levy, R., & Lo, W. J. (2015). Type I and Type II error rates and overall accuracy of the revised parallel analysis method for determining the number of factors. *Educational and Psychological Measurement, 75,* 428–457. http://dx.doi.org/10.1177/0013164414546566

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika, 19,* 149–161. http://dx.doi.org/10.1007/BF02289162

Harshman, R. A., & Reddon, J. R. (1983). Determining the number of factors by comparing real with random data: A serious flaw and some possible corrections. *Proceedings of the Classification Society of North America at Philadelphia, 14,* 14–15.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7,* 191–205. http://dx.doi.org/10.1177/1094428104263675

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

Humphreys, L. G., & Ilgen, D. R. (1969). Note on a criterion for the number of common factors. *Educational and Psychological Measurement, 29,* 571–578. http://dx.doi.org/10.1177/001316446902900303

Humphreys, L. G., & Montanelli, R. G., Jr. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research, 10,* 193–205. http://dx.doi.org/10.1207/s15327906mbr1002_5

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

Knol, D. L., & Berger, M. P. F. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research, 26,* 457–477. http://dx.doi.org/10.1207/s15327906mbr2603_5

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin, 109,* 502–511. http://dx.doi.org/10.1037/0033-2909.109.3.502

McDonald, R. P., & Ahlawat, K. S. (1974). Difficulty factors in binary data. *British Journal of Mathematical & Statistical Psychology, 27,* 82–99. http://dx.doi.org/10.1111/j.2044-8317.1974.tb00530.x

Mulaik, S. A. (2010). *Foundations of factor analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall.

Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research, 14,* 485–500.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swifts's electric factor analysis machine. *Understanding Statistics, 2,* 13–43. http://dx.doi.org/10.1207/S15328031US0201_02

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

R Core Team. (2015). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment, 12,* 287–297. http://dx.doi.org/10.1037/1040-3590.12.3.287

Ruscio, J., & Kaczetow, W. (2008). Simulating multivariate nonnormal data using an iterative algorithm. *Multivariate Behavioral Research, 43,* 355–381. http://dx.doi.org/10.1080/00273170802285693

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

Saccenti, E., & Timmerman, M. E. (2017). Considering Horn's parallel analysis from a random matrix theory point of view. *Psychometrika, 82,* 186–209. http://dx.doi.org/10.1007/s11336-016-9515-z

Shohat, J. (1929). Inequalities for moments of frequency functions and for various statistical constants. *Biometrika, 21,* 361–375.

ten Berge, J. M., & Kiers, H. A. (1991). A numerical approach to the approximate and the exact minimum rank of a covariance matrix. *Psychometrika, 56,* 309–315. http://dx.doi.org/10.1007/BF02294464

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16,* 209–220. http://dx.doi.org/10.1037/a0023353

Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement, 58,* 541–568. http://dx.doi.org/10.1177/0013164498058004001

Vale, C. D., & Maurelli, V. A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika, 48,* 465–471.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In E. Helmes (Ed.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41–71). New York, NY: Kluwer Academic/Plenum Press. http://dx.doi.org/10.1007/978-1-4615-4397-8_3

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442. http://dx.doi.org/10.1037/0033-2909.99.3.432

Received November 23, 2017
Revision received April 12, 2019
Accepted April 22, 2019 ■