Spring 2018

The Influence of a Proposed Margin Criterion on the Accuracy of Parallel Analysis in Conditions Engendering Underextraction

Justin M. Jones
Western Kentucky University, justinmjones23@gmail.com

THE INFLUENCE OF A PROPOSED MARGIN CRITERION ON THE ACCURACY OF PARALLEL ANALYSIS IN CONDITIONS ENGENDERING UNDEREXTRACTION

A Master's Thesis
Presented to
The Faculty of the Department of Psychological Sciences
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Justin M. Jones

May 2018

THE INFLUENCE OF A PROPOSED MARGIN CRITERION ON THE ACCURACY OF PARALLEL ANALYSIS IN CONDITIONS ENGENDERING UNDEREXTRACTION

Date Recommended 4/6/2018

Reagan D. Brown, Director of Thesis

Elizabeth L. Shoenfelt

Andrew Mienaltowski

Dean, Graduate Studies and Research 4/24/18 Date

ACKNOWLEDGEMENTS

I would like to first thank Dr. Reagan Brown for his instruction and mentorship throughout this program generally and this project specifically. You have helped develop my passion for research, nurtured my statistical skills, and have been a constant source of much needed levity. To Dr. Betsy Shoenfelt, my success on this project, and in this program, would not be possible without you. I thank you for always providing me with opportunities to learn and apply myself, and for sharing your wisdom with me on numerous occasions. I would also like to thank Dr. Mienaltowski, whose advice was invaluable throughout this entire project. I am also especially grateful to my family, who have always supported me in pursuit of my dreams. None of my success would be possible without your support. And finally, to my girlfriend Sydney, who is someone I can always count on, my biggest fan and supporter, and the person who always believes in me even when I don't. You drive me to achieve the best version of myself every day.

Table of Contents

INTRODUCTION.....1

Principal Component Analysis 1

Factor Analysis 2

Parallel Analysis 11

The Present Study 20

METHOD 22

Population Generation 22

Procedure 23

RESULTS 24

Sample Size Comparisons.....24

Iteration Comparisons 25

Percentile Comparisons 25

Margin Comparisons.....26

DISCUSSION 29

Discussion of Findings.....29

Limitations 30

Future Research 31

Conclusion 32

REFERENCES.....33

List of Tables and Figures

Table 1: Population Correlation Matrix.....22

Table 2: Accuracy Percentages of Factor Determination Criteria: Absolute Margin.....27

Table 3: Accuracy Percentages of Factor Determination Criteria: 10% Margin.....28

THE INFLUENCE OF A PROPOSED MARGIN CRITERION ON THE ACCURACY OF PARALLEL ANALYSIS IN CONDITIONS ENGENDERING UNDEREXTRACTION

Justin M. Jones May 2018 38 Pages

Directed by: Reagan D. Brown, Elizabeth L. Shoenfelt, and Andrew Mienaltowski

Department of Psychological Sciences Western Kentucky University

One of the most important decisions to make when performing an exploratory factor or principal component analysis regards the number of factors to retain. Parallel analysis is considered to be the best course of action in these circumstances as it consistently outperforms other factor extraction methods (Zwick & Velicer, 1986). Even so, parallel analysis could benefit from further research and refinement to improve its accuracy. Characteristics such as factor loadings, correlations between factors, and number of variables per factor all have been shown to adversely impact the effectiveness of parallel analysis as a means of identifying the number of factors (Pearson, Mundfrom, & Piccone, 2013). Critically, even the choice of criteria on which to evaluate factors (such as the eigenvalue at the 50^{th} or 95^{th} percentile) can have deleterious effects on the number of factors extracted (Peres-Neto, Jackson, & Somers, 2004). One area of parallel analysis yet to be researched is the magnitude of the difference between the actual eigenvalue and the random data-based eigenvalue. Currently, even if the margin between the actual eigenvalue and the random data-based eigenvalue is nominal, the factor is considered to be meaningful. As such, it may behoove researchers to enforce a higher standard, such as a greater margin between the two eigenvalues than just a simple difference. Accordingly, the purpose of this study was to evaluate the efficacy of a 10% margin criterion as compared to an absolute margin. These margins were evaluated in

conjunction with the 50th, 90th, 95th, and 99th percentile eigenvalue criteria on a population correlation matrix designed to engender underextraction. Previous research (Matsumoto & Brown, 2017) explored the same conditions on a population correlation matrix designed to elicit overextraction. They found that the most stringent standard (99th percentile eigenvalue plus a 10% margin) was the most accurate. For the present study however, it was hypothesized that the most accurate results would be obtained from a standard less stringent than the 99th percentile eigenvalue plus a 10% margin. The results suggest that when a correlation matrix has properties which may illicit underextraction, the use of less stringent criteria may lead to greater accuracy in identifying the number of factors and that the incorporation of an additional margin criterion may not improve the accuracy of the analysis.

Introduction

Being able to accurately and reliably determine the dimensionality of data is of great importance to researchers and practitioners alike. By identifying the dimensionality of a dataset, one is able to parsimoniously explain vast amounts of data using a relatively small number of factors or components without significant loss of information (Hoyle & Duvall, 2004). The most common statistical procedures to perform this kind of dimension reduction are factor analysis (common factor model) and principal component analysis (PCA). While functionally similar, PCA and factor analysis are based on different assumptions of measurement. As a result, each method is situated to answer a categorically different set of questions. Fortunately, this means that PCA and factor analysis are widely applicable as data reduction techniques across a variety of disciplines such as psychology, education, sociology, public health, management, economics, finance, ecology, chemistry, and even genomics (Dobriban, 2017).

Principal Component Analysis

Although many erroneously consider PCA a method of factor analysis, PCA is actually a data reduction technique (Gorsuch, 1990). PCA differs from genuine factor analysis in a number of important ways. The first difference is a conceptual one. In the common factor model of factor analysis, it is assumed that scores on the variables are caused by the latent constructs and error (Conway & Huffcutt, 2003). This relationship is what allows for the results to be interpreted in a causal manner and for inferences to be made about the function of the latent constructs. PCA however, defines factors, called components, based on scores on the manifest variables. With this measurement model

there is no identified cause of the scores and, therefore, the scores cannot provide any information regarding the latent constructs (Ford, MacCallum, & Tait, 1986).

Another important difference between the common factor model and PCA is that the common factor model factors the reduced correlation matrix which contains communalities along the diagonal. These communalities are the amount of variance which is shared amongst the variables, with the unique variance removed (Costello & Osborne, 2005). Conceptually, this means that the common factor model assumes that variables are measured with error. In contrast, PCA factors all of the variance, meaning that the communalities among factors are 1.0. The significance is that PCA assumes that all variance is common and that there is no error variance (Velicer & Jackson, 1990). As such, the interpretations that can be made about the results of a PCA are limited. Simply put, PCA can be seen as a method of identifying orthogonal axes which explain as much of the variance in the data as possible, using the least number of axes in order to do so.

Factor Analysis

Factor analysis is a family of multivariate statistical analyses which aim to reduce the number of variables while still retaining as much of the original variance accounted for by those variables as possible (Cokluk & Kocak, 2016). Broadly, factor analysis consists of two types of analyses: exploratory and confirmatory. Exploratory factor analysis is used when trying to identify the latent factor structure of a correlation matrix for which there is no theory or hypotheses on the number of factors which are present (Conway & Huffcutt, 2003). Confirmatory factor analysis however, is used on correlation matrices for which there is previous theory or the number of factors has been previously identified (Costello & Osborne, 2005). For this reason, confirmatory factor analysis is

best used to provide evidence for a theorized relationship amongst variables (Hurley et al., 1997).

Exploratory Factor Analysis

As mentioned, the purpose of exploratory factor analysis is to reduce a set of variables while also identifying the latent factor structure. Exploratory factor analysis is most commonly used to test construct validity, evaluate measures, and conduct hypothesis testing (Conway & Huffcutt, 2003). The most prevalent form of exploratory factor analysis is the common factor model. In the common factor model, variance is broken down into two parts: common variance, variance shared among the variables, and unique variance, variance not shared among the variables. Unique variance is further decomposed into systematic variance and random error (Costello & Osborne, 2005). Random error is unique variance in that it does not correlate with other variables and thus is specific to the variable. Systematic variance occurs when a variable measures a construct or variable which the other variables do not (Conway & Huffcutt, 2003). The most important kind of variance for the common factor model is common variance.

Common variance, henceforth referred to as communality, is the percent of variance a given variable shares with the other variables (Fabrigar, Wegener, MacCallum, & Strahan, 1999). However, this variance cannot be calculated and must be estimated (Hoyle & Duvall, 2004). The most common method to estimate communality is to compute the squared multiple correlation between a given variable and the other variables (Humphreys & Ilgen, 1969). Each variable is regressed on every other variable and an R^2 value is determined. These communality estimates are then placed on the diagonal of what is called the reduced correlation matrix. When the factor analysis is performed, the

communality estimates are further broken down into factor loadings. The factor loadings are estimated by optimally distributing the estimated communality from each variable to the factors based on the strength of their relationship (Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005). Thus, the factor loadings indicate how much variance in the given variable is accounted for by a given factor. The factor loadings from all of the variables are then used to determine eigenvalues.

Eigenvalues represent the amount of variance a factor accounts for in the dataset. As such, stronger factors will have larger eigenvalues while weaker factors will have smaller eigenvalues respectively. Strong factors are generally those which have multiple variables which load strongly on them (Hayton, Allen, & Scarpello, 2004). The purpose of exploratory factor analysis then, is to create a model consisting of meaningful factors while excluding trivial, or noise, factors. In other terms, the goal is to identify factors which account for a significant amount of variance in the dataset such that the factor structure will be sufficiently explanatory but also replicable across various datasets.

Issues with Factor Analysis

Using the correct method of analysis. For measure development, identifying latent factor structure, and other situations where error of measurement naturally occurs, factor analysis should be used instead of PCA as PCA extracts irrelevant variance from the items (Steger, 2010). For example, if the true data has low communalities, the PCA model will still assume communalities are 1.0. This would lead to great disparity between results of a PCA and the common factor model. Alternatively, if the communality between variables is high, the results of the two methods would likely be very similar (Ford et al., 1986).

Rotation. An important decision to be made when performing a factor analysis is what kind of rotation to use. Factor rotation is a procedure which adjusts the factor loadings in order to reveal simple structure (Conway & Huffcutt, 2003). Simple structure is when variables load highly on one factor and have low loadings on all other factors (Fabrigar et al., 1999). This structure is preferred because it makes the factors readily identifiable, and demonstrates the clarity of the factor solution. Unfortunately, there is not one form of factor rotation, nor is there only one rotated solution. In fact, there are an infinite number of ways which the factor solution can be represented in multi-dimensional space (Fabrigar et al., 1999). The method of rotation simply chooses the optimal solution based on the specified parameters.

There are two general categories of rotations: orthogonal and oblique. Orthogonal rotations maintain the uncorrelated factor structure that exists in an unrotated solution. Although orthogonal rotations provide solutions which are easy to interpret, an issue that arises is that generally factors are correlated to some degree. Thus, forcing factors which are correlated to be uncorrelated negatively influences the solution (Wood, Tataryn, & Gorsuch, 1996). Alternatively, oblique rotations allow factors to correlate. The issue with oblique rotations, however, is that the results are difficult to interpret and it is computationally more intensive (Wood et al., 1996). When deciding on what factor rotation to use, the main consideration should be the natural correlation between the factors. Unless there is reason to believe that the factors are uncorrelated, an oblique rotation should be used (Ford et al., 1986). While determining the method of analysis to use and type of rotation is important, arguably the most important decision that must be made is how to determine the correct number of factors.

Identifying the correct number of factors. Identifying the correct number of factors is more important than one may think. There are numerous consequences associated with overextraction (identifying too many factors) and underextraction (identifying too few factors) (Wood et al., 1996). For example, underextraction can lead to leaving important factors out of a model and the inclusion of significantly more error in the model (Wood et al., 1996), and can even lead to the accidental combination of distinct factors (Hayton et al., 2004). Importantly, if the number of factors is underidentified then variables which load on factors that have been left outside of the model will falsely load on factors included in the model, resulting in distorted factor loadings (Hayton et al., 2004). Although generally not considered as severe, overextraction also has negative consequences. For example, overextraction can lead to the retention of minor factors in the model of little or no significance (Fabrigar et al., 1999). These additional factors can then reduce the amount of variance available for variables to load on the legitimate factors. Over factoring can even result in factors that have no variables which load on them, or a singular large loading, leading to overly complex solutions and theories (Fava & Velicer, 2010). More generally, the impact of not determining the right number of factors is that the results will not replicate, making it difficult to establish claims of construct or structural validity (Steger, 2010). Therefore, determining the correct number of factors is of the utmost importance to someone conducting a factor analysis. There are numerous techniques which are used to determine the number of factors such as scree plot, eigenvalue-greater-than-one rule, parallel analysis, minimum average partial correlation and Bartlett's test.

Scree plot. One of the most common methods of determining the number of factors to retain is the scree plot. In order to determine the number of factors using a scree plot, first a graph is formed by plotting the eigenvalues of all of the factors. The slope between the eigenvalues is then examined. Where there is significant slope, is where there are factors which explain significant amounts of the available variance. The nominal factors are located where the slope is not as strong. As such, when examining a scree plot, one can determine the number of factors by determining the point where the slope flattens out and counting the number of factors to the left of that point (Costello & Osborne, 2005). One glaring issue with the scree plot method however, is that determining the number of factors requires a subjective opinion (Hayton et al., 2004). As a result, two researchers could view the same scree plot and determine two different factor solutions. Corrections to this issue have been postulated, such as having multiple raters evaluate the scree plot and calculating interrater agreement, however the subjectivity of the analysis is still an issue (Hayton et al., 2004). Another related issue with the scree plot approach is that sometimes there are no obvious changes in slope, or there are multiple drastic changes in slope, making it difficult to identify the number of factors (Turner, 1998). In these circumstances, the scree plot can be highly inaccurate and trivial. For these reasons, the scree plot is recommended for use only as a method for validating a factor solution which has been identified using a more accurate method (Fabrigar et al., 1999).

Eigenvalue greater than one. The eigenvalue greater than one rule, also referred to as the Kieser rule, Kieser's greater than one rule, or $K > 1$, is another relatively easy to implement rule for determining the number of factors to retain from a factor analysis or

PCA. The theoretical basis for the eigenvalue greater than one rule is that in a population matrix of uncorrelated data, the eigenvalues of components in a PCA are expected to equal one (Dinno, 2009b). This is because in PCA the components are standardized, which means that each variable accounts for exactly one unit of variance (Zwick & Velicer, 1986). Therefore, an eigenvalue of 1.0 functions as the lower bound of possible values that a component may take (Dinno, 2009b). As such, the eigenvalue greater than one rule states that if a component has an eigenvalue greater than 1.0 then it must be meaningful because it is, ostensibly, accounting for variance from other variables (Cliff, 1988). It is important to note that this rule may also be used to determine the number of factors when performing a factor analysis, but instead of a lower bound of 1.0, which represents the standardized variance of a component, the lower bound of uncorrelated factors is zero.

The eigenvalue greater than one rule may be the most popular method of determining the number of factors to retain. This, however, is troublesome because of its numerous limitations. First, almost all factor or principal component analyses are conducted on sample matrices. These sample matrices may have initial factors or components with eigenvalues greater than the lower bound solely due to sampling error and capitalization of chance (Cliff, 1988). Second, this method raises the question of what is to be made of marginal differences. For instance, in a PCA, is a component with an eigenvalue of .99 less significant than a component with an eigenvalue of 1.01?

Minimum average partial correlation. Minimum Average Partial Correlation (MAP) is a method of identifying the correct number of factors by analyzing the partial correlations within the reduced correlation matrix (Ruscio & Roche, 2012). The process

consists of first partialing out a factor from the matrix and then recalculating the average partial correlation between the remaining factors. Factors continue to be partialed out sequentially and average partial correlations are calculated after each factor is partialed out. This process continues as long as the average partial correlation continues to decrease. When the average partial correlation reaches its minimum, the number of factors that have been partialed out at that point are retained (Ruscio & Roche, 2012).

The rationale behind this method is that as common variance is removed from the model, the MAP criterion will decrease. However, once there is only unique variance left in the model, the MAP criterion will begin to increase. If one examines the equation for calculating a partial correlation, this relationship may be observed. In particular, the only circumstance where the partial correlation may increase is when the denominator decreases faster than the numerator. Such an event would occur, in these circumstances, only if a component or factor had a strong correlation with one variable and a weak correlation with all other variables, signifying a unique component or factor (Velicer, Eaton, & Fava, 2000). Significantly, the MAP method of identifying the correct number of factors to retain has been shown to be generally accurate, and applicable to any kind of covariance matrix (Zwick & Velicer, 1986). Zwick and Velicer found, in a test of five different methods of determining the number of factors to retain, that MAP was the second most accurate method. Their conclusions about MAP were that although it is accurate, it has a tendency to underestimate and may ignore smaller major components (Zwick & Velicer, 1986). Other studies have suggested that MAP is negatively biased and significantly influenced by correlations between factors (Pearson et al., 2013), and is most effective for small sample sizes (Zorić & Opačić, 2013).

Bartlett's Test. Unlike the other methods of determining the correct number of factors to extract from a factor analysis, Bartlett's test is a hypothesis test (Zwick & Velicer, 1986). Bartlett's test evaluates the null-hypothesis that all of the remaining eigenvalues in the model are equal starting with the first factor. If the null is rejected the eigenvalue of the next factor is tested. This continues until the test fails to reject the null. Then, all of the factors that were tested before the failure to reject the null are retained in the model (Hayton et al., 2004). Critically, it has been found that the accuracy of Bartlett's test is highly variable. For example, Zwick and Velicer (1986) found that factors such as sample size, alpha level, and the presence of unique variables, all had a significant impact on the accuracy of Bartlett's test.

Having reviewed many of the most common methods for determining the number of factors, it should be clear that there has been a need for an accurate and reliable method. Many factor identification methods are either subjective, inaccurate, or greatly influenced by fluctuations in the parameters of data such as factor loadings, sample size, number of variables and number of factors. Thus, it is suggested to use different methods in different scenarios, and to use multiple methods in order to verify solutions. While this kind of strategy may be effective, it adds even more decisions to an already very complex statistical analysis. Further, the decisions made during a factor analysis ultimately can influence the final outcome (Fabrigar et al., 1999). Adding even more decisions can negatively impact accuracy and replicability of factor solutions. In order to address many of these concerns, Horn (1965) developed a procedure called parallel analysis, which has shown great promise in the field of psychometrics.

Parallel Analysis

Parallel analysis has been shown repeatedly to be the most accurate method for correctly identifying the number of factors in a dataset (Velicer et al., 2000). Moreover, parallel analysis outperforms other methods across variations in sample size (Green, Thompson, Levy, & Lo, 2015), factor loadings (Buja & Eyuboglu, 1992), number of variables per factor (Crawford et al., 2010), and even distributional forms of data (Dinno, 2009a). Parallel analysis originated from Horn (1965), who developed the procedure as an improvement upon the eigenvalue greater than one rule. Horn argued that, although a population matrix of uncorrelated data will have eigenvalues of 1.0 (in a principal components analysis), a sample of data taken from that population will have eigenvalues greater than 1.0 for at least one factor simply due to chance. As such, it is a mistake to assume, as the eigenvalue greater than one rule suggests, that any eigenvalue which is greater than one represents a legitimate component (Franklin, Gibson, Robertson, Pohlman, & Fralish, 1995). Horn suggested that modeling the influence of sampling error on the size of eigenvalues could prove to be a more rational approach to factor identification. Thus, parallel analysis was developed with the intention of taking sampling error into account in order to more accurately identify legitimate factors (Horn, 1965).

Parallel analysis addresses the issue of sampling error by comparing the eigenvalues obtained from an analysis of real data to eigenvalues obtained from an analysis of simulated data. To conduct a parallel analysis, a population matrix is generated with the same parameters as the correlation matrix under analysis. Specifically, the simulated matrix has the same sample size and number of variables. A certain number

of sample matrices are then generated from this population matrix and are subsequently factor analyzed. The eigenvalues taken from the random sample matrices are then turned into a frequency distribution of eigenvalues for each factor. Simultaneously, the real correlation matrix is subjected to a factor analysis. The eigenvalues for each factor from the real data are then compared to the median (50^{th} percentile) eigenvalue from the respective random data eigenvalue frequency distribution. If the eigenvalue of the real factor is greater than the 50^{th} percentile eigenvalue of the random data, then the factor is retained. If the eigenvalue of the real factor is less than the 50^{th} percentile eigenvalue of the random data, then the factor is not retained, and no more factors are extracted (Ford et al., 1986).

Issues with Parallel Analysis

Even though parallel analysis is suggested to be the most accurate method of identifying the correct number of factors from a factor analysis (Hayton et al., 2004; Weng, 2005; Zwick & Velicer, 1986), there are still numerous variables which can influence its accuracy. Some of the most important of these influences are sample size, the number of factors, factor loadings, the number of variables, correlations between factors, and the interdependence of eigenvalues.

Sample size. The first issue with parallel analysis is that of sample size. Although sample size is a factor analysis issue, it also impacts the accuracy of parallel analysis. The accuracy of parallel analysis has been shown in various situations to decrease as sample size decreases (Crawford et al., 2010; Green et al., 2015). Zwick and Velicer (1986) found that, regardless of the number of variables, as sample size decreased the accuracy of parallel analysis also decreased. Zorić and Opačić (2013) found similar results in a

study comparing five component retention criteria, including parallel analysis, across variations in sample size, the number of variables, the number of components, and the percentage of error variance. The authors found that, all else being equal, at small sample sizes (i.e., $N = 50$) parallel analysis was less accurate than other methods such as MAP. Much of this variability in performance due to sample size can be attributed to sampling theory; as sample sizes increase the standard deviation of the sampling distributions decrease. As sample size increases, there is inherently less variation in the sampling distribution which allows for more accurate comparisons between the simulated eigenvalues and the real data eigenvalues.

Sample size also influences parallel analysis accuracy through its effects on eigenvalue size. For example, Turner (1998) found that sample size may significantly influence the size of noise eigenvalues following the identification of a real factor. In a study of the size of noise eigenvalues across variations in sample size and percent of common variance, Turner found that at low sample sizes (i.e., $N = 100$), as the amount of common variance increased the size of subsequent eigenvalues decreased sharply. However, for large sample sizes (i.e., $N = 1000$), as the amount of common variance increased the subsequent eigenvalues decreased only slightly (1998). Importantly, traditional parallel analysis does not control for the amount of variance accounted for by the preceding factors. Considering this fact within the context of the Turner (1998) study, traditional parallel analysis may be prone to underextraction especially at small sample sizes.

Number of variables. Parallel analysis has also been suggested to be influenced by the number of variables. For example, Green, Thompson, Levy, and Lo (2012)

performed a Monte Carlo study of parallel analysis accuracy under various conditions. The authors were able to demonstrate that as the number of variables per factor increased from four to eight, the accuracy of parallel analysis decreased. Crawford et al. (2010) found similar results in their Monte Carlo study of different parallel analysis criteria. In the Crawford et al. study, the authors demonstrated that when comparing three different eigenvalue criteria, the conventional 50th percentile criterion became less accurate as the number of variables per factor increased from three variables per factor to six. Further, in a study investigating the accuracy of five different factor retention criteria across variations in sample size (i.e., 50, 100, 200, 300, or 600), number of factors (i.e., 1, 2, 3, 5, 8, or 10), number of variables (i.e., 9, 15, 22, 35, or 40), and amount of error variance (i.e., 30, 40, 50, 60, 70, or 80), Zorić and Opačić (2013) found that as the number of variables increased, the propensity for parallel analysis to underextract increased. The results of their study also suggested that as the factor to variable ratio increases, accuracy of parallel analysis decreases. One possible explanation for such behavior is that well defined factors usually consist of a relatively small number of variables with high loadings on their respective factor and low loadings on the other factors (Zwick & Velicer, 1986). Having a large number of variables per factor leads to poorly defined factors which are harder to identify through parallel analysis.

Number of factors. Another influence on the accuracy of parallel analysis is the number of factors present. Although in reality a scientist or practitioner will not know how many factors are present, especially when performing an exploratory factor analysis, research has studied matrices of known factor structure. In these studies, such as that by Crawford et al. (2010), the authors have suggested that the accuracy of parallel analysis is

negatively influenced by the number of factors present such that parallel analysis will be less accurate in identifying the correct number of factors when there are more factors in the dataset. For example, in a study by Beauducel (2001) which investigated the accuracy of parallel analysis on orthogonal and oblique solutions, it was found that although increasing the number of factors had no impact on the accuracy of parallel analysis on orthogonal solutions, increasing the number of factors within oblique solutions was coupled with noticeable decreases in accuracy.

Factor loadings. The factor loadings, or how much variance in a variable is accounted for by a specific factor, may also influence the accuracy of parallel analysis (Crawford et al., 2010). Under ideal circumstances, the factors in a factor analysis are well defined. This means that there are relatively few variables defining each factor; these variables have strong loadings on the factor and have weak loadings on the other factors (Zwick & Velicer, 1986). In a study by Zwick and Velicer (1986), the authors evaluated the accuracy of numerous criteria for identifying the number of factors under different sample sizes (i.e., twice or five times the number of variables), factor loadings (i.e., .50 or .80), and numbers of variables (i.e., 36 or 72). Critically, their study suggested that the ability of parallel analysis to identify the correct number of factors increased as factor loadings increased. Crawford et al. (2010) also found evidence suggesting that the accuracy of parallel analysis may increase as factor loadings increase. However, it is not uncommon for variables to load moderately on numerous factors. These cross loaded variables make it notably more difficult to identify the correct number of factors.

Correlations between factors. One of the greatest influences on the accuracy of parallel analysis for determining the number of factors from a factor analysis is the

correlation between factors. In the broadest sense, factors correlate when variables in the dataset are correlated with other variables. Several studies have shown that high correlations among factors leads to inaccuracies (Crawford et al., 2010; Green et al., 2015). For example, in the Crawford et al. (2010) study of parallel analysis methods for determining the number of factors, the authors found that both PCA and principal axis factoring performed worse when there were high (i.e., .7) correlations among factors as well as small (i.e., .4) factor loadings. They also found that criteria, such as the 95th percentile eigenvalue criterion, had increased levels of underextraction as the correlations between factors increased (Crawford et al., 2010). It is suggested that the correlation between factors leads to underextraction because eigenvalues are interdependent (Hayton et al., 2004). Specifically, the size of an eigenvalue is determined by the size of other eigenvalues. Thus, if factors are highly correlated, the size of the eigenvalues will be limited, increasing the likelihood of underextraction (Turner, 1998).

Joint distributions. Another issue which has recently received increased attention is determining the correct reference distribution to be used for parallel analysis. Many argue that the size of the initial eigenvalue in traditional parallel analysis influences the size of subsequent eigenvalues (Buja & Eyuboglu, 1992; Green, Xu, & Thompson, 2017; Turner, 1998). This arises from the fact that each variable contributes a specific amount of variance. As such, the sum of the contributed variance from all variables is equal to the total variance which is analyzed. When one factor accounts for a significant proportion of this total variance, the eigenvalues of the subsequent factors will naturally decrease due to the limited amount of variance left to account for (Turner, 1998). Traditional parallel analysis treats each factor's eigenvalue distribution as a

completely separate analysis, without accounting for any influence that the previous factor may have had. Importantly, this can lead to underextraction as the random sample distribution will most likely have larger values than the eigenvalues of factors from real data which follow a large factor (Turner, 1998). While this assumption is appropriate for the first eigenvalue, for subsequent eigenvalues the traditional method which assumes that there are no factors present is inappropriate.

In order to correct for this, revised forms of parallel analysis have been suggested (Green et al., 2015; Green et al., 2017). These forms of parallel analysis operate under the assumption that all of the eigenvalues are interdependent. Thus, once the first factor has been identified, the subsequent eigenvalue distribution takes into account this factors presence, reducing the amount of variance available to provide a more accurate estimation for comparison (Green et al., 2015). These methods have been found to be somewhat more accurate than the traditional method under certain conditions. For example, Green et al. (2015) found that revised parallel analysis performed better than traditional parallel analysis when there were highly correlated factors, weak factor loadings, or a high number of variables per factor. Significantly, they also found that while the accuracy of traditional parallel analysis decreased with increases in sample size, the accuracy of revised parallel analysis increased.

Criteria used for factor identification in parallel analysis. The final issue related to parallel analysis to be discussed is that of the eigenvalue criteria that are used. In Horn's (1965) seminal article he proposed that the eigenvalue from the real data be compared to the median (i.e., 50th percentile) eigenvalue of the random data. It has been suggested, however, that this criterion can lead to overextraction, particularly through the

retention of minor components (Buja & Eyuboglu, 1992). In response, more stringent criteria have been proposed such as the use of the 95th or 99th percentile eigenvalue. The 95th percentile criterion was first suggested by Glorfeld (1995) who reasoned that Horn's (1965) parallel analysis tended to extract too many or poorly defined factors. He postulated that implementation of a more stringent criterion could correct for this tendency for parallel analysis to overextract.

In many ways, the 95th percentile criterion emulates hypothesis testing (Weng, 2005). As such, the 95th percentile criterion has been regarded as an effective heuristic, though not a statistically rigorous model (Saccenti & Timmerman, 2017). For example, an eigenvalue is compared to a frequency distribution and if this eigenvalue is greater than the 95th percentile eigenvalue then the factor is retained. In this instance we accept that there is a 5% chance that the factor is actually the result of sampling error. With this logic, it is clear as to why the 95th percentile criterion might be preferred. Using the same logic, when considering the 50th percentile eigenvalue we are accepting that there is a 50% chance that the factor is due to sampling error. Fundamentally this means that the median eigenvalue criterion makes it more likely that factors which are due to sampling error are retained (Green et al., 2015).

There have been a number of studies which have compared the median eigenvalue criterion to the 95th and even 99th percentile eigenvalue criteria. For example, Glorfeld (1995) was able to demonstrate in his seminal study that, when factors were poorly defined, the median eigenvalue led to overextraction, while this was not so when using the 95th or 99th percentile criteria. Crawford et al. (2010) compared the median and 95th percentile eigenvalue criteria, as well as principal axis factoring and PCA, across various

factor loadings, numbers of factors, and sample sizes. The authors found that the 95th percentile criterion outperformed the median criterion as sample size and items per factor increased, as well as when the factor loadings were medium to low. However, when there were small sample sizes, high correlations between factors, or few variables per factor, the median eigenvalue criterion performed better than the 95th percentile eigenvalue criterion (Crawford et al., 2010). Green et al. (2015) also compared the 95th percentile criterion against the median eigenvalue criterion by testing them on different sample sizes, numbers of factors, factor loadings, and factor correlations. The authors found that the 95th percentile criterion was generally most effective across all conditions; however, accuracy was impacted by the definition of the factors. Further, in a test of parallel analysis on unidimensional binary data, Weng (2005) found that the 95th, and even more stringent 99th percentile eigenvalue criteria, were almost always 100% accurate and greatly outperformed the median eigenvalue criterion.

Although there have been several findings in support of using more stringent criteria (Cokluk & Kocak, 2016; Horn, 1965; Peres-Neto et al., 2004), troublesome issues remain with using criteria such as the 95th percentile eigenvalue. For example, Turner (1998) cautions that using this criterion may lead to overestimation of the size of noise eigenvalues. Noise eigenvalues correspond with factors which are analyzed after a real factor has been identified. Turner argues that this is especially true for traditional parallel analysis which does not treat the eigenvalues as interrelated. Also, studies such as that by Steger (2010), have shown that the 95th percentile eigenvalue criterion still is influenced by factors such as high commonalities and correlations between factors. In summation, even though parallel analysis combined with the 95th percentile eigenvalue criterion

seems to be the most accurate method of identifying the correct number of factors to date, ways to make this method more robust while also increasing its accuracy are needed.

The Present Study

The present study investigated the efficacy of an additional criterion to be used in conjunction with eigenvalue percentile criteria. This additional criterion demands that an eigenvalue from the true data be greater than the specified percentile eigenvalue from the random distribution by a certain margin. For this study, the performance of a 10% margin was compared against the traditional simple difference standard of parallel analysis. Further, these standards (any difference vs 10% difference) were combined with four different eigenvalue criteria (50^{th}, 90^{th}, 95^{th}, 99^{th}), yielding eight different standards for determining the number of factors with a parallel analysis. Previous research (Matsumoto & Brown, 2017) investigated the efficacy of a margin criterion in conjunction with eigenvalue criteria on a correlation matrix designed to engender overextraction. The authors found that the most stringent criteria (i.e., 99^{th} percentile eigenvalue combined with 10% margin) was more accurate in identifying the number of factors in comparison to all other criteria. However, this rule will not likely be the most effective standard in correlation matrices designed to engender underextraction.

In correlation matrices which engender underextraction, generally, there is a strong correlation among factors. This reduces the probability of large eigenvalues as variance is distributed from the variables across the factors. Peres-Neto et al. (2004) demonstrated this phenomenon in an investigation of 10 stopping rules on principal component analysis. Of the methods tested, the authors examined the accuracy of parallel analysis in conjunction with the 95^{th} percentile criterion and 50^{th} percentile criterion. The

authors found that, when the components were highly correlated, the median eigenvalue outperformed the 95th percentile criterion. The authors suggested that a criterion which was greater than the median, but less than the 95th percentile would maximize the probability of identifying the correct number of non-trivial factors when correlations between components were present while subsequently reducing the number of trivial components retained in the presence of uncorrelated variance (Peres-Neto et al., 2004). Therefore, it is hypothesized that less stringent criteria than those used by Matsumoto and Brown (2017) will be the most accurate. In particular, it is postulated that a slightly less stringent percentile criterion (i.e., 90th percentile) with an absolute margin criterion will yield the correct number of factors more often than will other standards. The reason for this hypothesis is that the correlation between factors should reduce the size of the eigenvalues in the real data, necessitating less stringent criteria to prevent underfactoring. *Hypothesis:* A 90th percentile eigenvalue criterion in addition to an absolute margin criterion will identify the correct number of factors at a higher rate than will any other combination of criteria.

The study consisted of a 3 (Sample Size) x 2 (Number of Iterations) factorial design with 2 (Margin Criteria) x 4 (Eigenvalue Criteria) decision rules investigated for each condition. The accuracy of these decision rules was tested using a Monte Carlo simulation. This simulation tested the various criteria using a population correlation matrix which was designed to elicit underextraction. In order to facilitate underextraction, the population correlation matrix was created to have a high correlation between factors. This effect was further tested upon variations in sample size and number of iterations of

the parallel analysis. The accuracy of the various criteria, determined by their ability to identify the predefined factor structure, was then compared.

Method

Population Generation

A dataset, consisting of 1,000,000 cases with scores on 12 variables was generated in order to create a population correlation matrix (Table 1) designed to engender underextraction in samples. The correlation matrix consisted of a two factor structure. Within this matrix, each of the factors was defined by five variables. The correlation between factors was .45. The remaining two variables were uncorrelated with all variables.

Table 1.

Population Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1	1.0	.35	.35	.35	.35	.25	.25	.25	.25	.25	.00	.00
2		1.0	.35	.35	.35	.25	.25	.25	.25	.25	.00	.00
3			1.0	.35	.35	.25	.25	.25	.25	.25	.00	.00
4				1.0	.35	.25	.25	.25	.25	.25	.00	.00
5					1.0	.25	.25	.25	.25	.25	.00	.00
6						1.0	.35	.35	.35	.35	.00	.00
7							1.0	.35	.35	.35	.00	.00
8								1.0	.35	.35	.00	.00
9									1.0	.35	.00	.00
10										1.0	.00	.00
11											1.0	.00
12												1.0

Procedure

The experiment adhered to the following procedure:

1. A sample of 120, 240, or 480 cases was randomly drawn from the population.
2. An exploratory factor analysis (common factor model) was performed on the sample data.
3. A parallel analysis of 200 or 500 replications was conducted on a sample of random data of the same size and with the same number of variables as the sample data.
4. The eigenvalues from the 50^{th}, 90^{th}, 95^{th}, or 99^{th} percentile from the parallel analysis were compared to the eigenvalues obtained from the factor analysis of the sample data.
5. The number of factors was then determined by applying a margin criterion (e.g., simple absolute difference or 10%). The number of factors was defined as the highest factor in the sample data which had a positive eigenvalue greater than the corresponding random parallel analysis eigenvalue. In the event that the eigenvalue of a lower factor from the sample data was not greater than the eigenvalue from the random parallel analysis data, but the next sample factor was, then all factors after the factor which was not greater than the random parallel analysis eigenvalue were ignored.
6. The process was repeated 1000 times at which point the results were compared against the known population value of two factors.

.

Results

The purpose of the current study was to investigate the accuracy of various eigenvalue percentile criteria when used in conjunction with a margin criterion on a population matrix designed to engender underextraction in samples. In order to evaluate the accuracy of the decision rules, a frequency analysis was performed to determine the percentage of iterations where the criteria correctly identified, over-identified, or under-identified, the predefined factor structure (i.e., two factors). These percentages, across variations in sample size and number of iterations, are presented below (Table 2 displays the results of the percentile criteria in conjunction with an absolute margin; Table 3 presents the accuracy of the percentile criteria in conjunction with a 10% margin). Due to the number of comparisons in the present study, only the most relevant results will be discussed.

Sample Size Comparisons

At the smallest sample size (N = 120), all criteria performed relatively poorly, with the best results identifying the correct number of factors in approximately 50% of the cases. The sample of (N = 120) also led to significant amounts of underextraction for any standard greater than the 50^{th} percentile. Less stringent criteria seemed to be the most accurate under these conditions, with the 50^{th} percentile criterion being generally more accurate at small sample sizes regardless of margin. Given the generally poor accuracy of any standard at this sample size, no further attention will be given to the results of this condition.

With a sample size of (N = 240), there was a noticeable increase in the accuracy of all criteria. Interestingly, at this sample size, there was still a substantial amount of

underextraction across the various decision rules. The best results were found with the 90th and 95th percentile criteria, with the 90th percentile performing only marginally better than the 95th. Even so, the highest level of accuracy was only 83%.

At the largest sample size (*N* = 480), there was again a noticeable increase in the accuracy of the criteria. Notably, the 99th, 95th, and 90th percentile criteria yielded very accurate results (correct factor identification rates greater than 95%), with the 95th and 99th slightly outperforming the 90th with an absolute margin, and the 95th and 90th performing slightly more accurately with a 10% margin. With a sample size of (*N* = 480) there was very little, if any, underextraction. Generally, the results show that larger sample size is better regardless of standard used with excellent success rates (> 90%) for any percentile in the 90s (i.e., 99th, 95th, 90th) with or without a 10% margin

Iteration Comparisons

Regardless of condition, the impact of the number of iterations on the accuracy of the various decision rules was trivial. As such, it is safe to conclude that no more than 200 iterations are necessary for parallel analyses matching the conditions of this study.

Percentile Criteria Comparisons

Across all conditions of the study, less stringent criteria (50th, 90th) generally outperformed the most stringent (95th, 99th) criteria regardless of margin both in terms of accuracy and percentage of underextraction. Only at large sample sizes (*N* = 480) did the most stringent criteria perform similar to, or better than, the less stringent criteria. In all but the largest sample size (*N* = 480), the 99th percentile was the least accurate at identifying the number of factors. Further, in every condition the 99th percentile led to the highest rates of underextraction.

Margin Criteria Comparisons

For the (N = 240) sample size, the absolute margin criterion in combination with the 99th, 95th, and 90th percentile criteria was the most accurate. The only percentile criterion used in conjunction with a 10% margin criterion to perform similarly to, or better than, the criteria with an absolute margin at this sample size was the 50th percentile rule. At the largest sample size (N = 480), the 10% margin in combination with the percentile criteria was as accurate as the absolute margin for the 90th, 95th, and 99th percentiles and was more accurate at the 50th percentile. A notable pattern here was that as sample size increased the difference in performance between the two margin criteria decreased significantly. At the largest sample size (N = 480), this effect was manifested through nominal differences in performance between the two margin criteria across all of the percentile criteria.

Table 2.

Accuracy Percentages of Percentile Criteria in Conjunction with an Absolute Margin

Iterations	Sample Size	Absolute Margin											
		99th			95th			90th			50th		
		Under	Correct	Over	Under	Correct	Over	Under	Correct	Over	Under	Correct	Over
200	120	69.6%	30.2%	0.2%	53.8%	44.4%	1.8%	45.2%	50.2%	4.6%	18.6%	56.3%	25.1%
500	120	73.5%	25.9%	0.6%	57.4%	40.2%	2.4%	50.1%	45.6%	4.3%	20.6%	54.4%	25.0%
200	240	29.7%	69.2%	1.1%	17.1%	80.3%	2.6%	12.3%	83.3%	4.4%	2.6%	75.8%	21.6%
500	240	29.0%	70.2%	0.8%	17.1%	80.6%	2.3%	11.7%	83.5%	4.8%	2.2%	72.4%	25.4%
200	480	1.3%	98.2%	0.5%	0.3%	98.3%	1.4%	0.2%	96.6%	3.2%	0.0%	83.6%	16.4%
500	480	0.7%	98.8%	0.5%	0.1%	98.2%	1.7%	0.0%	96.7%	3.3%	0.0%	84.9%	15.1%

Note: Percentages reflect the number of iterations out of 1000 where the specified rule identified the same number of factors as in the population matrix (correct), fewer factors than in the population matrix (under), or more factors than were present in the population matrix (over).

Table 3.

Accuracy Percentages of Percentile Criteria in Conjunction with a 10% Margin

Iterations	Sample Size	10% Margin											
		99th			95th			90th			50th		
		Under	Correct	Over	Under	Correct	Over	Under	Correct	Over	Under	Correct	Over
200	120	84.5%	15.4%	0.1%	72.7%	27.1%	0.2%	63.0%	36.1%	0.9%	32.4%	55.4%	12.2%
500	120	85.6%	14.4%	0.0%	73.9%	25.6%	0.5%	66.7%	31.8%	1.5%	35.3%	52.1%	15.6%
200	240	48.5%	51.2%	0.3%	31.3%	67.5%	1.2%	22.7%	75.6%	1.7%	5.8%	83.6%	10.6%
500	240	47.0%	53.0%	0.0%	30.4%	68.7%	0.9%	22.0%	76.6%	1.4%	5.6%	81.9%	12.5%
200	480	2.8%	97.1%	0.1%	1.4%	98.2%	0.4%	0.7%	98.6%	0.7%	0.1%	92.1%	7.8%
500	480	3.1%	96.8%	0.1%	0.8%	98.9%	0.3%	0.4%	98.7%	0.9%	0.0%	92.7%	7.3%

Note: Percentages reflect the number of iterations out of 1000 where the specified rule identified the same number of factors as in the population matrix (correct), fewer factors than in the population matrix (under), or more factors than were present in the population matrix (over).

Discussion

This study investigated the influence of a correlation matrix designed to elicit underextraction on the accuracy of various parallel analysis eigenvalue decision rules when used in conjunction with a margin criterion. It was hypothesized that because the correlation matrix was designed to elicit underextraction, a less stringent criterion (i.e., 90th percentile criterion with absolute margin) would be the most accurate across conditions. Our results partially support this hypothesis. Although the 90th percentile criterion with an absolute margin was not the most accurate decision rule across all conditions, less stringent criteria (i.e., 90th, 50th) were generally more accurate. In Matsumoto and Brown's (2017) study on the accuracy of parallel analysis criteria on a matrix designed to elicit overextraction, they found the most stringent criterion (i.e., 99th percentile with a 10% margin) to be the most accurate. The results of our study support our hypothesis that an alternative, less stringent criterion would be most accurate in identifying the number of factors in a correlation matrix engendering underextraction. Importantly, this demonstrates that the most effective decision rule for data which engender overextraction is not necessarily the most accurate for data which engender underextraction. Such findings emphasize the fact that there is not a singular universal decision rule to be used in all scenarios. More specifically, it seems imperative for those performing parallel analyses to determine the number of factors to take the intricacies of their dataset into consideration when deciding on a criterion to use.

Our results support the findings of previous research on a number of different aspects. In particular, we were able to demonstrate, much like Glorfeld (1995) that, in general, a more stringent criterion than Horn's original 50th percentile rule is needed to

improve the accuracy of factor identification. Further, these results reiterate the findings of Peres-Neto et al. (2004) and Crawford et al. (2010), who found that less stringent criteria were more accurate in the presence of correlated factors. Peres-Neto et al. (2004) suggested that a criterion that was more stringent than the 50th percentile yet less stringent than the 95th percentile (such as the 80th percentile) would be the most effective in the presence of highly correlated variance. Our results lend support to this suggestion. Our findings also coincide with those of Green et al. (2012) who found that less stringent criteria are more accurate at low sample sizes (i.e., N = 100) and, in particular, when there are two highly correlated factors or low factor loadings.

Limitations

The main limitation of the present study was that traditional parallel analysis was used. As mentioned by Turner (1998), a significant issue with traditional parallel analysis is that the distributions of eigenvalues following a legitimate factor are not adjusted according to the variance accounted for by that factor. In other words, a previously identified factor will account for a significant amount of the available variance thus reducing the amount of variance that can be assigned to subsequent factors. That being said, traditional parallel analysis treats the test of each factor as unrelated. Critically, this effect can lead to underextraction under certain circumstances (Buja & Eyuboglu, 1992). There are revised forms of parallel analysis available which use the appropriate reference distribution, and studies have been conducted demonstrating the efficacy of these methods (Green et al., 2012; Green et al., 2017); however, they are still relatively nascent and computationally intensive. In spite of this theoretical limitation, the results in some conditions indicated factor identification levels so high (correct number of factors

identified over 95% of the time) that any improvement offered by modified parallel analysis would be minimal in those situations.

Another potential limitation is that only one population matrix was studied. It is very possible that with matrices containing different numbers of factors or different correlations between factors, outcomes could have been substantially different. To partially address this limitation, the population matrix used for the study was designed to model many of the common issues faced when performing parallel analyses. Other matrices should be subsequently explored.

Future Research

Reflecting on the current findings, there are a number of areas that future research should investigate. One such area is the accuracy of margin criterion in conjunction with the various percentile criteria under different conditions such as number of factors, correlations between factors, and numbers of variables per factor. Such research could further support the results of this study, or identify boundary conditions which explain the behavior of these decision rules. Future research should also seek to investigate the accuracy of these rules using the revised method of parallel analysis. Very little research has investigated these relatively novel methods, so developing a greater understanding of their behavior and or influences is of great importance. Finally, the results suggest that in conditions engendering underextraction, a criterion which is greater than the 50^{th} percentile but less than the 90^{th} percentile might be the most effective. Future research should investigate other percentile standards which may be used in order to maximize accuracy in identifying the number of factors under these conditions.

Conclusion

Through a Monte Carlo study, we were able to demonstrate that, under conditions which illicit underextraction, a less stringent criterion is generally more accurate for identifying the number of factors when performing a parallel analysis. The present study also adds to the body of research demonstrating that there is not one universal decision rule which can be used in all situations. Rather, our results suggest that one should be cognizant of the intricacies of their dataset when deciding on the decision rule to use when performing a parallel analysis. Finally, the results of our study suggest that parallel analysis performs poorly at small sample sizes (10:1 subjects to variables ratio), with large amounts of underextraction, regardless of criteria. Alternatively, at large sample sizes (40:1 subjects to variables ratio), all factor identification criteria achieved optimal performance (i.e., > 90% factor identification), except for the 50^{th} percentile criterion. Furthermore, the additional use of a 10% margin does not seem to influence the accuracy of the percentile criteria at large sample sizes. Similarly, at medium sample sizes (20:1 subjects to variables ratio), our results suggest that moderate percentile criteria (i.e., 90^{th} and 95^{th}) without the use of a 10% margin are the most accurate. Overall, these results indicate that in conditions which engender underextraction, an additional 10% margin criterion does not seem to provide any additional utility.

References

Beauducel, A. (2001). Problems with parallel analysis in data sets with oblique simple structure. *Methods of Psychological Research Online*, *6*, 141–157. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/issue14/art2/article.html

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

Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin*, *103*, 276–279. http://doi.org/10.1037/0033-2909.103.2.276

Cokluk, O., & Kocak, D. (2016). Using Horn's parallel analysis method in exploratory factor analysis for determining the number of factors. *Educational Sciences: Theory & Practice*, *16*, 537-551. http://doi.org/10.12738/estp.2016.2.0328

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6*, 147–168. http://doi.org/10.1177/1094428103251541

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*, 1–9. http://doi.org/10.1.1.110.9154

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://doi.org/10.1177/0013164410379332

Dinno, A. (2009a). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research, 44*, 362–388. http://doi.org/10.1080/00273170902938969

Dinno, A. (2009b). Implementing Horn's parallel analysis for principal component analysis and factor analysis. *Stata Journal*, *9*, 291–298. Retrieved from https://www.stata-journal.com/article.html?article=st0166

Dobriban, E. (2017, October 2). Factor selection by permutation. Retrieved from https://arxiv.org/abs/1710.00479

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://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://doi.org/10.1207/s15327906mbr2703_5

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*, 291–314. http://doi.org/10.1111/j.1744-6570.1986.tb00583.x

Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T., & Fralish, J. S. (1995). Parallel analysis - a method for determining significant principal components. *Journal of Vegetation Science*, *6*, 99–106. http://doi.org/10.2307/3236261

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://doi.org/10.1177/0013164495055003002

Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research, 25,* 33-39. http://doi.org/10.1207/s15327906mbr2501_3

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://doi.org/10.1177/0013164411422252

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://doi.org/10.1177/0013164414546566

Green, S. B., Xu, Y., & Thompson, M. S. (2017). Relative accuracy of two modified parallel analysis methods that use the proper reference distribution. *Educational and Psychological Measurement,* 1-16. http://doi.org/10.1177/0013164417718610

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://doi.org/10.1177/1094428104263675

Hogarty, K., Hines, C., Kromrey, J., Ferron, J., & Mumford, K. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement, 65,* 202-226. http://doi.org/10.1177/0013164404267287

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

Hoyle, R. H. & Duvall, J. (2004). Determining the number of factors in exploratory and confirmatory factor analysis. In D. Kaplan (Ed.). *The SAGE Handbook of Quantitative Methodology for Social Sciences* (pp. 301-315). Thousand Oaks, CA: Sage.

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://doi.org/10.1177/001316446902900303

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior, 18,* 667–683. http://doi.org/10.1002/(SICI)1099-1379(199711)18:6<667::AID-JOB874>3.0.CO;2-T

Matsumoto, M. & Brown, R. D. (2017). An investigation of the accuracy of parallel analysis for determining the number of factors in a factor analysis. *Honors College Capstone Experience/Thesis Projects*. Paper 696. http://digitalcommons.wku.edu/stu_hon_theses/696

Pearson, R., Mundfrom, D., & Piccone, A. (2013). A comparison of ten methods for determining the number of factors in exploratory factor analysis. *Multiple Linear Regression Viewpoints, 39,* 1-15. Retrieved from http://www.glmj.org/archives/articles/Pearson_v39n1.pdf

Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2004). How many principal components? Stopping rules for determining the number of non-trivial axes

revisited. *Computational Statistics & Data Analysis*, *49*, 974-997. http://doi.org/10.1016/j.csda.2004.06.015

Ruscio, J., & Roche, B. (2011). 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://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://doi.org/10.1007/s11336-016-9515-z

Steger, M. F. (2010). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*, *86*, 263–272. http://doi.org/10.1207/s15327752jpa8603

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://doi.org/10.1177/0013164498058004001

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 *Problems and Solutions in Human Assessment* (Vol. 1998, pp. 41–71). http://doi.org/10.1007/978-1-4615-4397-8_3

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some further observations. *Multivariate Behavioral Research*, *25*, 1-28. http://doi.org/10.1207/s15327906mbr2501_12

Weng, L.-J., & Cheng, C. P. (2005). Parallel analysis with unidimensional binary data. *Educational and Psychological Measurement*, *65*, 697–716. http://doi.org/10.1177/0013164404273941

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, *1*, 354–365. http://doi.org/10.1037/1082-989X.1.4.354

Zorić, A., & Opačić, G. (2013). Impact of different conditions on accuracy of five rules for principal components retention. *Psihologija*, *46*, 331–347. http://doi.org/10.2298/PSI130801008Z

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432. http://doi.org/10.1037/0033-2909.99.3.43