Journal of
EDUCATIONAL IMPACT

Examining Retention Methods in Factor Analysis: A Comparison of Polychoric and Pearson Correlations for Categorical and Continuous Data

Sebahattin Ziyanak[1]*, Jason Yagci[2]

Abstract: *In the last two decade, it has become clear that retention methods must utilize polychoric correlation instead of Pearson correlation to eliminate drawbacks such as underestimation of the magnitude of the relationship between latent variables that result in spurious findings (Bernstein & Teng, 1989). In the present study, the literature review will be examined and compared using Monte Carlo simulation to determine the most parsimonious method of retention for categorical data. With continuous variables, the majority of researchers still implement Cattell's scree test (Henson & Roberts, 2006) and Kaiser-Guttman-1 rule (Velicer et al., 2000), because these procedures are the default in popular statistical packages, such as SPSS and SAS. The present study will examine two of the most accurate methods: MAP (Minimum Average Partial) and PA (Parallel Analysis) along with Very Simple Structure (VSS) with categorical variables.*

Keywords: Monte Carlo Simulation, Cattell's Scree Test, Kaiser-Guttman-1 Rule, Minimum Average Partial, Parallel Analysis, Very Simple Structure

Introduction

One of the main challenges in measurement instrument development for educational purposes is the need to identify latent variables based upon observed variables. Generally, because relationships exist between latent variables, most instruments encompass more than one trait. Factor analysis is one of the methods utilized to identify the latent traits that instruments measure. Factor analysis is the general term for two techniques in literature: exploratory factor analysis (EFA) and principal component analysis (PCA). Since EFA extracts more factors empirically than the actual number of true factors that exist, the decision of the number of factors to retain becomes a crucial decision for analysts. Many different methods and procedures have been introduced and developed for factor retention, but the majority of these methods were originally designed for continuous variables. Moreover, especially in education and psychology, instrumentation consists

[1] University of Texas Permian Basin, Associate Professor of Sociology, E-mail: ziyanak_s@utpb.edu

[2] Independent Researcher

of dichotomous and categorical item indicators therefore creating a need for the examination of retention methods and techniques that are optimal for categorical variables.

The advancement in computer technology and software has led to less time in computation for factor analysis, especially for some retention methods that utilize simulation techniques. Even though the availability of software and the speed of computation make it possible to use multiple methods to determine the number of factors to retain, the K-G-1 rule is still the most popular retention method since it is the default method in some statistical packages.

Due to the more complex computations required for common factor analysis; principal component analysis (PCA) has been utilized with the majority of extraction methods. These methods employ the Pearson correlation, which is the most effective type of correlation for interval and ratio scale data (Zwick & Velicer, 1986; O'Connor, 2000). However, many instruments in education and psychology use Likert scale items, which are ordinal. As such, the polychoric correlation should be implemented to eliminate biases such as the underestimation of the strength of the relationship between factors.

Most comparison studies of factor retention focused on instruments with items measured on ratio and interval scales. As a result, Pearson correlation was used to determine factor loadings (Zwick & Velicer, 1986; Fabrigar et al., 1999). Systematically, categorical and ordinal scale data, which are less descriptive than interval and ratio scale data, are understudied in factor retention studies. Though several studies (Timmerman & Lorenzo-Seva, 2011; Garrido et al., 2013; Garrido et al., 2011) have separately examined retention methods, MAP and PA, with ordinal and categorical data, none has compared the methods to determine which is optimal with categorical data.

Definition of Terms

Communality: "The squared factor loading represents the proportion of variance in the indicator that is explained by the latent factor" (Brown, 2006, p. 61).

Eigenvalues: "Area-world, variance-accounted-for statistics that characterize the amount of information present in a given factor or function. (Eigenvalues are also sometimes synonymously called "characteristic roots)" (Thompson, 2004, p. 178).

Factor: "A composite variable, which consists of the loading or correlation between that factor and each variable making up that factor. Factor analysis is used to determine the extent to which a number of related variables can be grouped together into a smaller number of factors which summarize the linear relationship between those variables" (Cramer & Howitt, 2004, p. 63).

Factor loading: "are completely standardized estimates of the regression slopes for predicting the indicators from the latent factor, and thus are interpreted along the lines of standardized regression (β) or correlation (r) coefficients as in multiple regression/correlational analysis" (Brown, 2006, p. 15).

Factor rotation: "Graphic visual or mathematical movement of the axes measuring the factor space used in exploratory factor analysis so that the factors can be more readily interpreted" (Thompson, 2004, p. 178).

Monte Carlo methods: "means of calculating, among other things, the probability of outcomes based on a random process. So any statistical test which is based on calculating the probability of a variety of outcomes consequent of randomly allocating a set of scores is a Monte Carlo method" (Cramer & Howitt, 2004, p. 104).

Oblique rotation: "a form of rotating factors in which the factors are allowed to intercorrelate (i.e., permit factor axis orientations of less than 90°). The correlation between two factors is equal to the cosine of the angle between the rotational axes. Because cos(90) = 0, the factors are uncorrelated in orthogonal rotation. In oblique rotations, the angle of the axis is allowed to be greater or less than 90°, and thus the cosine of the angle may yield a factor correlation between 0 and 1.0" (Brown, 2006, p. 31).

Orthogonal rotation: "the process by which the factors are constrained to be uncorrelated (i.e., factors are oriented at 90° angles in multidimensional space). In applied social sciences research, orthogonal rotation is used most often, perhaps because it is the default in major statistical programs such as SPSS (varimax rotation)" (Brown, 2006, p. 31).

Principal axis factoring: "a form of factor analysis in which only the variance shared between the variables is analyzed. Variance, which is unique to a variable or is error, is not analyzed. The shared variance or communality can vary from a minimum of 0 to a maximum of 1. It is generally less than 1." (Cramer & Howitt, 2004, p. 130).

Principal components analysis: "aims to account for the variance in the observed measures rather than explain the correlations among them. Thus, PCA is more appropriately used as a data reduction technique to reduce a larger set of measures to a smaller, more manageable number of composite variables to use in subsequent analyses " (Brown, 2006, p. 22).

Literature Review

In this literature review chapter, factor analysis is broadly described. In addition the similarities and differences between principal components and common factor analysis are discussed. The next section of this chapter includes descriptions of several methods for determining the number of factors to retain, such as Horn's parallel analysis, Velicer's minimum average partial method, and Revelle's very simple structure. In the last section, comparisons of these retention methods with Pearson and polychoric correlations are explained.

Factor Analysis

Even though factor analysis has been utilized for more than one hundred years, new analytic methods in combination with more powerful statistical programs and computers have permitted researchers from social sciences to health sciences to analyze data faster and more accurately. In addition, factor analysis has been utilized as a tool for evaluating construct validity of psychological instruments. When Spearman (1904) introduced factor analysis more than a century ago, he examined the interrelationships of students' abilities in different subject areas along with the identification of observed characteristics; he related all these in a common factor, which he named "general intelligence. Cudeck and MacCallum (2007) described factor analysis as "one of the great success stories of statistics in social science." This success was attributed to the technique's ability to deal with popular subjects such as intelligence, social class, and health status, in which there are many unobservable variables.

Hayton et al., (2004) pointed out that the decision the researcher makes regarding the number of factors to retain becomes more crucial in EFA than Confirmatory Factor Analysis (CFA) due to the fact that EFA is widely used in scale development even with a weak theoretical background. The importance of this decision making is summarized (Hayton et al., 2004) as follows; first, decision of the number of factors to retain has more impact on the results of the factor analysis than choice of methods and or type of rotation. Second, the balance between the number of factors to retain and the variance explained by the factors should be maintained in order to differentiate the major and minor factors. Third, specifying fewer or many factors would affect the results by leading to poor factor loading pattern production and interpretation.

Fabrigar et al., (1999) advise balance between the minimum number of factors retained and the adequacy of this number to explain and cover the model. Additionally, the goal should be to determine the number of factors that describe the data given relevant theory, bearing in mind that this decision comes with some selection errors, which have a direct effect on the result. One such crucial error is under-factoring or over-factoring.

The result of under-factoring or over-factoring is well documented (Fava & Velicer, 1996; Wood et al., 1996; Zwick & Velicer, 1986). Studies suggest that under-factoring is less desirable because it is more likely to lead to substantial error. If indicators that are supposed to be extracted as a separate factor are not included in the model but loaded on to other extracted factors, this results in poor estimation of factor loadings. Likewise, if two factors are combined and rotated for a single common factor, the result can be difficult to interpret because of the complex pattern of factor loadings (Fabrigar et al., 1999). Research on the number of factors to retain indicates that over-factoring produces fewer problems in factor loading estimation than under-factoring. In over-factoring, the extra factor would likely have one or more measured variables from each of the major factors. This would lead into the next step in PCA, where the extra factor merges as the major factor with the help of larger factor loadings (Fabrigar et al., 1999).

Factor analysis is categorized into two groups according to the purpose of the usage of the procedure: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). If the number of factors or components were already known according to theory, CFA would be the proper procedure. On the other hand, if the number of factors is unknown and the study is examining a new instrument, then EFA is the appropriate procedure. In the present study, EFA will be the focal point due to the nature of the procedure. A critical initial step for researchers is determination of the number of factors to retain when conducting an EFA. Generally, in statistical packages, K-G-1 rule, PA, and MAP methods are designed to handle continuous variables, such as interval and ratio scales. However, there are vast amounts of instruments that utilize ordinal items in social science, especially in education.

Floyd and Widaman (1995) categorized procedures for determining the number of factors to retain into three groups: statistical tests, mathematical and psychometric criteria, and rules of thumb. Maximum likelihood and the generalized least squares are statistical tests (Floyd & Widaman, 1995). These statistical tests are no longer acclaimed for defining the number of factors to retain (Velicer et al., 2000). The Kaiser-Guttman-1 rule (Kaiser, 1960), parallel analysis (Horn, 1965), and minimum average partial (Velicer, 1976) are mathematical and psychometric criteria. Cattell's scree test (Cattell, 1966), percentage of variance accounted for (Stevens, 2002), and number of variables that have significant loading for a factor are rules of thumb (Gorsuch, 1983). Several independent researchers in the last three decades showed that mathematical and psychometric criteria are most frequently utilized by researchers (Velicer et al., 2000) while others are implemented as default methods in statistical software packages (such as, Kaiser- Guttman-1). Zwick and Velicer (1986) and Fabrigar et al., (1999) indicated that Velicer's Minimum Average Partial method (Velicer, 1976) and Horn's Parallel Analysis (Horn, 1965) are the most dependable and accurate methods for retention decisions in factor analysis with continuous variables even though they are not implemented as one of the default methods. Although Velicer's Minimum Average Partial method (Velicer, 1976) and Horn's Parallel Analysis (Horn, 1965) are well studied with items measured on a continuous scale, it is unclear how well existing retention procedures perform with ordinal items. Even when the methods that have demonstrated efficacy are used to retain the factors, retention decisions need to be interpretable and theoretically practical. While the investigation is exploratory by nature, theory and previous research should provide a frame of reference for the optimum number of factors to retain (Fabrigar et al., 1999).

It is important in exploratory factor analysis that the intent of the measure is directly related to observed variables; consequently, the sample of observations selected should be careful and thoughtful to lead to optimum results with EFA (Henson & Roberts, 2006). In other words, the factor can be identified as a cluster of several variables that are a homogenous set of items measuring similar traits. Under this description, the purpose of factor analysis is to explain a larger set of measured variables with a smaller set of latent variables (Henson & Roberts, 2006). Under a narrower perspective, EFA is utilized as a tool for identifying factor structure. In other words, the preference is for a minimum number of factors to explain as much as possible about the variance in observed variables.

Interpretability of factors has been noted as an important part of EFA investigations by several researchers (Worthington & Whittaker, 2006). If interpretable and plausible factors are retained, factor analysis would be considered to reach the researcher's goals. Even though strong empirical evidence may exist to support a factor, the retention of factors should depend on conceptual interpretability (Worthington & Whittaker, 2006). A significant amount of research has compared PA, MAP and VSS (retention methods) with continuous variables. On the other hand, because very few studies have examined them with ordinal items, there is need for comparison studies in this area.

Comparison of Principal Component and Common Factor Analysis

Principal Component Analysis (PCA) is the default technique of extraction in current statistical packages (Costello & Osborne, 2005). Even though PCA and factor analysis are utilized to reduce a larger number of variables to a smaller number of factors, both procedures have their own unique steps. PCA is the only extraction method that can be calculated by hand. Fabrigar et al., (1999) pointed out that principal axis factoring is recommended when data is not normally distributed. Moreover, significant difference between PCA and common factor analysis occurs when factors are poorly identified with low saturation (Velicer & Jackson, 1990).

Comparison of Exploratory Factor Analysis and Confirmatory Factor Analysis

If little theoretical knowledge is available to identify factors, two main issues remain in EFA (Liu & Rijmen, 2008), they are: (a) what is the number of factors to retain, and (b) what is the relationship between observed and latent variables. This would be helpful in enabling model development that more accurately represents the data in EFA. It is unlikely that there are ever a finite number of factors resulting in a *true model*, even though many analysts proceed with this idea in mind. However, using a theoretical model-selection approach (Preacher et al., 2013) may be more helpful to identify a model that best balances model-data fit, where models are selected from a set of competing theoretical explanations from observed variables and attempt to develop the most parsimonious representation of factors that are worthwhile to retain, versus a fallacious *correct* number of factors approach. Generally, analysts have very little or no prior knowledge for selecting the reasonable number of factors in EFA. However a theoretical background that supports model-based decision-making will result in a better solution with factor analysis.

EFA is a "data-driven" approach (Brown, 2006) because often the number of factors is unknown at the beginning of the analysis as well as relationships between variables and factors. EFA may also be used for confirmatory purposes (Gorsuch, 1997) because EFA can be conceptualized as a two-tailed test, whereas confirmatory factor analysis, under this same conceptualization, can be thought of as a one-tailed test. In scale development, the early stage of exploratory factor analysis is therefore an opportunity to test the dimensionality of the scale. EFA can be helpful to identify, describe, and filter factors. During initial steps in scale development where EFA is appropriate, an iterative process of item deletion plays an important role because factors are affected by deleted items due to changes in factor loadings and cross-loadings on factors (Worthington & Whittaker, 2006).

Methods for Determining the Number of Factors to Retain

Although Kaiser- Guttman-1 and Cattell's scree test are default methods in popular statistic packages, the two methods most often recommended for factor retention: Horn's parallel analysis (PA), and Velicer's minimum average partial method (MAP) will be compared in the present study. In addition, a relatively new method for factor retention: Revelle's very simple structure (VSS) will be compared in the present study.

Horn's Parallel Analysis

Horn's parallel analysis is one of the methods used for deciding the number of factors to retain in principal components and principal factor analysis (Horn, 1965). Even though it was not a popular analytic technique among social science researchers in the last three decades of the 21st century, it has proven its accuracy in testing the dimensionality of factors (Zwick & Velicer, 1986; Velicer et al., 2000). In addition, the availability of written code for SAS and SPSS (O'Connor, 2000), made this method more frequently applied in the last decade. Since the Kaiser-Gutmann-1 rule utilizes population statistics and does not consider or include sampling error and least squares bias, the PA method (Horn, 1965), which estimates the proportion of variance due to sampling error, is recommended for use in factor analysis.

Horn (1965) takes some further steps from Kaiser's (1960) proposed retention method which retain the components with eigenvalues greater than one, and formulates his technique with finite number of observations as follows:

First, conduct a parallel PCA on an n by p matrix of random values. Second, repeat this k times; and third, average the eigenvalues λ_q^r over k, to produce $\overline{\lambda_q^r}$ and lastly, adjust λ_q by subtracting from it (λ_q^r -1) to produce λ_q^{adj}(Dinno, 2011, p.4).

So, the retention technique can be summarized as in Equation 1 and 2 (Dinno, 2011, p.4):

$$\lambda_q^{adj} \begin{cases} > 1 \ retain \\ \leq 1 \ do \ not \ retain \end{cases} \quad (1)$$

$$\lambda_q \begin{cases} > \overline{\lambda_q^r} \ retain \\ \leq \overline{\lambda_q^r} \ do \ not \ retain \end{cases} \quad (2)$$

The PA procedure uses Monte Carlo simulation, which begins with sample size and the number of variables in an intended real data set. With these two parameters, the simulation generates matrices on random data. In PA the results involve comparing eigenvalues of the original sample correlation matrix and the generated random data matrix to decide the number of factors to retain (Buja & Eyuboglu, 1992). Parallel analysis generates correlation matrices where parameters are based on the sample size and the number of variables in empirically collected data (or *real* data). When the mean of eigenvalues from the generated correlation matrices are compared to eigenvalues from a real data correlation matrix, at some point observed eigenvalues become smaller than the generated eigenvalues, this crossover provides a cut point to determine the number of factors to retain.

There is no limitation in generating the number of matrices of random data; however, the same values must be used for the generating data; these include sample size and the number of variables in the actual data. Similar to K-G-1, where there is no practical difference between 1.0 and 0.99, investigators can make clear decisions with such miniscule differences using PA (Fabrigar et al., 1999).

Even though PA is a well-established method to determine the number of factors to retain, it tends to underestimate the number of factors when the first eigenvalue is relatively large compared to others (Beauducel, 2001). As shown by many studies on PA (Zwick & Velicer, 1986; Fabrigar etal., 1999; Peres-Neto et al., 2005;), the procedure is most accurate when factors are orthogonal and the sample is large. With smaller sample sizes, and less simple structure, PA is likely to under-extract with oblique data (Beauducel, 2001). The accuracy of PA is directly related to sample size and inversely related to the number of components presented in the population matrix. Under-extraction may be solved (to some extent) with a large sample size.

PA performs well with both Pearson and Polychoric correlations (Cho et al., 2009). Polychoric based PA produces fewer factors because of higher correlation among factors, and yields larger initial factors.

Several different versions of PA are reported. The most common PA procedure with PCA as described by Crawford et al. (2010) requires multiple sample correlation matrices be generated under assumptions of normally distributed and uncorrelated population data. It is also important that sample size and the number of variables are the same as observed data. A variation of PA is proposed by Buja and Eyupoglu (1992) where instead of generating random data with normality assumptions, the procedure requires generating simulated data using permutations of existing data. This version of PA does not produce significantly different results than the original version.

Increasing sample size, loadings, and the number of variables will improve the performance of PA (Crawford et al., 2010). When high correlations exist among factors and the average factor loading is smaller, results of PA tend to over or under extract factors. It has been shown that PA using polychoric correlation with ordinal data results in better accuracy for identifying the number of factors to retain (Timmerman & Lorenzo-Seva, 2011). But it is difficult to apply the polychoric procedure with empirical data because convergence is often unattained.

Horn (1965) developed PA based on common factor analysis; however, some researchers. (Buja & Eyuboglu, 1992; Crawford, et al., 2010) have pointed out that because of using an unreduced correlation matrix, PA is more appropriate for PCA. PA appears to work well with both common factor and PCA.

Velicer's Minimum Average Partial Method

Similar to PA, MAP (Velicer, 1976) has been difficult to conduct because it is not implemented in popular statistical packages (O'Connor, 2000). Several studies (Zewick & Velicer, 1986; Velicer et al., 2000) indicate that MAP has a tendency to under-extract the number of factors, and as already stated, under-extraction is perceived to be a more serious problem than over-extraction in factor analysis (Zewick & Velicer, 1986; Velicer et al., 2000). The matrix of partial correlation is central to the MAP retention method in the context of principal component analysis (Velicer, 1976). Velicer et al. (2000) briefly summarized the steps of MAP as follows: "First each component is partialed out of the correlation matrix and the partial correlation matrix is calculated. Second, for the number of variables, the average of the squared correlation matrix is computed. Lastly, the number of components to retain is indicated at the point where the average squared partial correlation reaches a minimum" (p. 44).

The matrix of partial correlations is obtained by first computing the partial covariance matrix (Velicer, 1976, p.322),

$$C_{11}^{*} = R - A\,A' \quad (3)$$

where C is the partial covariance matrix, R is the correlation matrix, and A is the pattern matrix. The partial correlation matrix is then computed

$$R^{*} = D'\,C\,D \quad (4)$$

where R* is the matrix of partial correlations and D is the diagonal of the partial covariance matrix. The MAP procedure involves determining when the matrix of partial correlations most approximates an identify matrix, i.e., determining the value m for which

$R^* = I$ (5)

Velicer (1976) proposes the statistic to determine the number of components to extract as follows:

$f_m = \sum \sum_{i \neq j} (r_{ij}^*)^2 / (p(p-1)$ (6)

where r_{ij}^* is the element in row *i* and column *j* of the matrix of partial correlations. The value of f_m is the average of the squared partial correlations after the first *m* components are partialed out. The proposed stopping point is the value of m for which f_m is at a minimum. The value f_m would be calculated from $m = 1$ to $p - 1$; the value of f_m for $m = p$ is indeterminate since the diagonal of C_{11}^* (partial covariance matrix) consist of zeros. The values of f_m will range between 0 and 1. A second summary of statistic, useful for comparative purposes, is

$f_0 = \sum \sum_{i \neq j} (r_{ij}^*)^2 / (p(p-1)$ (7)

If $f_1 > f_0$, then no components would be extracted (p. 323).

Garrido et al., (2011) summarized the findings of different versions of MAP as follows: “All different versions of MAP perform better with a larger sample size, higher factor loadings, more variables per factors, less number of factors, lower factor correlations, and smaller skewness” (p. 556). The other conclusion was that all versions of MAP procedures underestimate the number of factors to retain. Additionally, factor loadings and the number of variables per factor are the most influential variables that affect the accuracy of the MAP method. Fabrigar et al., (1999) point out that MAP is designed only for PCA and not recommended for common factor analysis. MAP is originally based on PCA; the rationale of the method is theoretically informed by CFA (Velicer, 1982, Velicer & Jackson, 1990). MAP standardizes residuals by converting them into partial correlations (Velicer, 2000). As mentioned earlier, a limitation of MAP is the tendency to under-factor when factor loadings are small and there are few observed variables per factor (Zwick & Velicer, 1986). Another drawback to MAP is that the cut off value is very close to the adjacent value; it is possible to get two lowest average square partial correlations which are very close to each other from MAP run. However, variations in sample size tend to produce minuscule effects on the accuracy of MAP (Velicer, 2000).

Revelle's Very Simple Structure

Very Simple Structure (VSS) utilizes comparisons of the goodness of fit of simple structure matrix with the initial correlation matrix to determine the optimal number of factors to retain from the correlation matrix (Revelle & Rocklin, 1979). VSS helps to answer not only the question of how many factors to retain but also the question of how to rotate the factors that have been retained (Revelle & Rocklin, 1979, p. 405). The general tendency in the interpretation of factor retention is that parallel to the nature of factor retention techniques, generally the largest loadings are examined while smaller loadings are ignored or under examined (Revelle, 2013). The default number of observation used with VSS is 1000, but it can be specified. VSS draw a plot to show model fits by means of number of factors (Revelle & Rocklin, 1979).

The steps of VSS are briefly listed by Revelle and Rocklin (1979, p 405-407) as follows: First, find an initial factor solution with k factors (principal factor can be utilized), and then in second step, rotate the solution to maximize the rotation criterion that is preferred, such as Varimax, or an oblique rotation. In the third step, the VSS criterion is applied in two step procedure; first of all, for the VSS solution of factor complexity v,replace the k-v smallest elements in each row of the factor pattern matrix with zero. This matrix is called simplified factor matrix S_{vk}. And then, to evaluate how well a particular rotated factor solution F_k fits a simple structure model of factor complexity v, consider how well the matrix:

$$R_v^* = S_{vk} \Phi S'_{vk} \quad (8)$$

(where Φ is the factor inter-correlation matrix) reproduces the initial correlations in R. That is, find the residual matrix:

$$\overline{R_v} = R - R_v^* = R - S_{vk} \Phi S'_{vk} \quad (9)$$

In the next step, as an index of fit of $\overline{R_v}$ to R, find one minus the ratio of the mean square residual correlation to the mean square original correlation:

$$VSS_{vk} = 1 - MS_r'/MS_r \quad (10)$$

where the degree of freedom for these mean square are the number of correlations estimated less the number of free parameters in S_{vk}. The mean squares are found for the lower off-diagonal elements in R and $\bar{R}$.

Finally, to determine the appropriate number of factors to extract, find the value of VSS criterion for all values of k from one to the rank of matrix. The optimal number of interpretable factors is the number of factors, k, which maximizes VSS_{vk} (Revelle & Rocklin, 1979, p.405-407).

Comparison of the Methods

Comparison of Methods with Pearson Correlation

A summary of study findings on factor retention methods based on Pearson correlation is presented in this section. Several studies propose that PA is the most suitable technique to decide the number of factors to retain (Humphreys & Montanelli, 1975; Glorfeld, 1995; Zwick & Velicer, 1986). Glorfeld (1995) also concluded, similar to Zwick and Velicer (1986) that PA is a better method compared to the other factor retention methods.

Minimum Average Partial Method (MAP; Velicer, 1976), similar to PA, utilizes principal component analysis and is established on partial correlations of variables. It also uses the EFA's common factors to determine how many factors or components to extract. In MAP the first step is to determine what components are common solution, without trying to find the cutoff point for the number of factors. MAP appears precise under many situations although under certain situations it may divulge a propensity to underestimate the number of factors (Zwick & Velicer, 1986). When there are small factor loadings and few variables per component, MAP consistently underestimates the number of major components (Zwick & Velicer, 1986).

Henson and Roberts (2006) recommended that multiple criteria be utilized to determine the number of factors to retain in EFA because this early decision will directly affect the rest of the analysis. Defining a factor and interpreting a factor depends on how many factors are retained; if over or under factoring occurs then conclusions leading to decisions about factors will be different because of the omitted or combined factors in the main factors (Henson et al. 2004). A study by Henson et al. (2004) showed that the K-G-1 rule is the most popular method, followed by the scree test. More interestingly, about one third of researchers utilize prior theory to aid in the number of factors to retain, which implies that CFA could be used instead of EFA. A large proportion of papers using EFA (22.4 %) failed to report the retention method that was used with the EFA and only 8 percent of the studies that performed EFAs used more than one retention rule which is recommended by many scholars (Henson et al., 2004). Two of the most accurate methods, PA and MAP were not used in any of the articles that were studied by Henson et al. (2004).

Another comparison study by Worthington and Whittaker (2006) summarized studies published in the Journal of Counseling Psychology between 1995 and 2004 using EFA, 18 percent of the studies used K-G-1 rule for factor retention decisions, and 17 percent utilized the scree test. No studies utilized the PA or MAP methods. Gorsuch (2003) acknowledged that although there are many retention methods, none of them are solely adequate. In EFA, there are several different variables that have various levels of contribution to the determination of the optimum number of factors to retain. For this reason, it is advisable to examine the results obtained from several different methods before determining the number of factors to retain by interpretability of the solution (Hayashi et al., 2007).

Velicer et al. (2000) study's conclusion could be compendious as follows: In terms of theoretical rationale, MAP is the most dependable method; the K-G-1 rule is the easiest rule to implement, yet the results of this method are exceptionally inaccurate and not recommended. PA is more accurate than MAP, although both are considered the most reliable methods. Velicer et al.,

(2000) drew a road map for researchers who need to determine the number of factors to retain in EFA as follows: Perform a component analysis at the early stages of EFA, use a combination of MAP and PA methods to determine the best solution, ensure that the final factor solution is interpretable and if not run the analysis a second time. The K-G-1 rule is not recommended and further should be excluded from statistical packages as the default method.

Thompson (2004) advises using several different strategies would help to make a decision easier especially if results confirm each other. He pointed out that because eigenvalues carry sampling error, a researcher should employ judgment when the K-G-1 rule is used to make a decision in factor retention. Henson and Roberts (2006) reported that 55 percent of published EFA studies utilized one criterion to determine the number of factors to retain. Therefore, they recommended that more than one method be used to check agreements; particularly in situations where there are a large number of factors expected to be retained. Five main elements, have been identified that would shape the results of the analysis in factor retention: (a) the size and composure of the sample, (b) selection of variables, (c) model fit, (d) rotation methods, and (e) the number of factors (Velicer et al., 2000).

Parallel with advancements in computer programming and availability of speed in computation to researchers in the last decade, MAP and PA were utilized more frequently as factor retention methods in item analysis. Both of these methods are based on principal component analysis that claims to be an approximation of CFA. Even though PA and MAP are the most accurate procedure for determining the number of factors to retain, these are not offered by popular statistical packages, but these procedures could be run with free software, such as R with the FACTOR package (Lorenzo-Seva & Ferrando, 2006). Fabrigar et al., (1999) advised using more than one method to examine the rotated solution for the model, and then compare the solution with theory and interpretability.

Comparison of Methods with Polychoric Variables

An EFA utilizing ordinal items and performed in SPSS showed that Velicer's MAP performed better when compared to other factor retention methods such as PA, VSS, scree test and K-G-1 (Basto & Pereira, 2012). However, K-G-1 rule overestimated the number of factors to retain, in spite of being the default procedure in some of the statistical packages. The typical correlation used in PA and MAP methods is the Pearson correlation because these methods depend on the linear relationship between factor scores and expected values. Yet this assumption may not be valid with ordinal items because the Pearson correlation would underestimate the relations between the items (Timmerman & Lorenzo-Seva, 2011).

In the last decade, several researchers (Cho et al., 2009; Garrido et al., 2013) implemented studies to investigate the analytic methods used to determine the number of factors to retain with ordinal items. These studies mainly focused on the two most accurate methods: MAP and PA. Cho et al. (2009) investigated PA with polychoric correlations, which are designed for ordinal items. The comparison of PA with Pearson correlation and PA with polychoric correlation indicates similar performances to determine the optimal number of factors. Polychoric correlation was shown to obtain an unbiased estimation of the relationship among observed categories.

Garrido et al., (2011) examined the performance of factor retention with categorical data using Velicer's Minimum Average Partial method. The investigation examined the relative impact of several variables on the accuracy of the MAP method with ordinal items. The manipulated items in this study were similar to the Cho et al., (2009) study, including factor loading, factor correlation, number of variables per factors, number of response categories, sample size, and skewness. There are three recommendations: (a) if the items are categorical, use polychoric correlations and the square of partial correlations instead of Pearson correlations, (b) smooth the non-Gramian polychoric matrices with the ridge procedure, and (c) design scales to be identified by at least six variables (Garrido et al., 2011). Garrido et al., (2011) specified that the most accurate MAP estimations resulted in the combination of these criteria: polychoric correlation instead of Pearson, use the instrument that has scales with at least six variables, and the mean of random eigenvalues instead of the 95th percentile.

Holgado–Tello et al., (2010) compared factor results with polychoric correlation and Pearson correlation, and concluded that when polychoric correlations were used for a measurement model, it produced a better result, regardless of the number of factors. Timmerman and Lorenzo-Seva (2011) also pointed out similar findings for polychoric correlation as more reliable relative to Pearson correlation for principle axis factoring in factor analysis.

Factor Retention for Categorical and Ordinal Data

In social science, most variables for factor analysis are items from scales that are measured with ordinal or nominal responses. While many different factor retention methods have been utilized in EFA with items measured on a continuous scale (Garrido et al., 2011), ordinal items have not been exposed to examination by researchers. Pearson correlation routinely underestimates the strength of relationship between categorical variables (Babakus et al., 1987; Bollen & Barb, 1981), or can produce spurious "difficulty" factors (Gorsuch, 1983). Consequently, Pearson correlation produces biased dimensionality estimates with categorical data.

Using ordinal items in factor analysis presumes that the relationship between the items is nonlinear (Cho et al., 2009). Basto and Pereira (2012) indicated that under PCA and factor analysis, Pearson correlation is the only option for such analyses therefore ordinal items are, by default, examined with Pearson correlation. Ordinal items can only be considered a rough representation of unobserved continuous items (Basto & Pereira, 2012).

Monte Carlo Simulation

In many situations it is impossible or impractical to obtain analytical data for statistical, and in particular psychometric analyses. However, it is possible to generate random data in the form of sampling distributions that conforms to prior established parameters such as the number of factors, factor intercorrelation, number of items, number of respondents through a computer simulation referred to as a Monte Carlo procedure. The general goal of Monte Carlo (MC) simulations is to reach maximum generalizability and replicability of obtained results. Studies often use Monte Carlo simulation to examine the effect of extreme potentials of variables, such as very large sample size, to identify the potential of a methodology; in the present study this would be a factor retention method (Hutchinson & Bandalos, 1997). Monte Carlo simulations can be programmed to generate sample covariance matrices with a specified population, correlation magnitude, number of items, and sample size (Choi et al., 2011). In other words, Monte Carlo simulation methods help researchers generate sample data by controlling and manipulating the variables to test the functionality of statistical analyses. Simulation also provides the opportunity to manipulate variables that represent a variety of possible samples.

Monte Carlo simulations provide researchers a tool to understand sampling distribution and random sampling (Mooney, 1997). It is advised empirically by MC simulation via "using random samples from a known population of simulated data to trade a statistics behavior" (Mooney, 1997, p.2).

Mooney (1997) explains the principles of Monte Carlo as follows: "the behavior of a statistic in random samples can be assessed by the empirical process of actually drawing lots of random samples and observing this behavior" (p.4). In addition, Mooney (1997) describes the complexities of generating Monte Carlo simulations and interpretation of estimated sampling distributions.

The number of replications implemented is crucial for MC simulation. Wilcox (1992) recommended 1000 replications for MC simulation, a conservative estimate (Wilcox, 1988). It should be noted however that a higher number of replications is preferred as this provides more stable estimates and more precise confidence intervals (Thoemmes et al., 2010). Mundform et al. (2011) recommended 8000 replications for the Monte Carlo simulation would be sufficient for stable results. MC simulations can help researchers to generate sampling distributions that are theoretical and unobserved (Paxton et al., 2001).

Hutchinson and Bandalos (1997) describe MC studies as data generation based on behavior of statistical estimators. MC studies' sampling distributions are fundamental to inferential statistics and mathematically generated according to known parameters of a population distribution, for example normality and sample size (Hutchinson & Bandalos, 1997). Similar to a typical research study, MC simulation involves several steps including identification of the population, description of sampling data, data collection, and data analysis (Hutchinson & Bandalos, 1997).

Hutchinson and Bandalos (1997) describe the steps of MC studies as follows: First, it is necessary to define the research question, as it shapes the analytic question. Second, the researcher needs to specify parameters, such as the number of replications and the design of the study. Third, the simulation program should be written to generate data in light of previous designs in the area. Finally, the generated sample data should be analyzed according to the design and hypothesis.

MC studies generate samples from distributional criteria, and then examine the totality of all sample distributions generated under imposed limits for variables. Outcomes represent probabilities of events that would occur under the specified conditions. These studies help the researcher to generate and control many variables, such as the size of a sample, which would be difficult to replicate in practice.

Harwell et al., (1996) advise that if a problem could not be solved analytically or mathematically, an MC study should be implemented as an alternative way of conducting the study. Moreover, other than using simulated data, MC studies are identical to empirical studies (Harwell et al., 1996). The parameters of variables have to be selected after variables to be manipulated are identified. In the first step of MC simulation, dependent variable(s) should be identified. Dependent variables must be specified with Type I and II error and/or permissible parameter estimate bias (Hutchinson & Bandalos, 1997). To run a Monte Carlo simulation, it is necessary to determine independent variables size and parameters (e.g. the range and or distributional shape of variables). The limitation of independent variables would affect the generalizability of the results (Hutchinson & Bandalos, 1997).

There are several advantages of using inferential statistics in MC simulation analysis: sampling error would be taken into account, and the magnitude of effect sizes would be estimated (Hutchinson & Bandalos, 1997). Harwell et al. (1996) pointed out that when an analytic solution does not exist or is unfeasible due to complexity, MC studies are essential for arriving at a solution. With MC studies, the combination of parameters are easily manipulated and specified. The simulation of large samples helps mitigate the cost of working with humans.

Researchers aim to generalize the pattern and result that are obtained from one sample of data to infer conclusions about the larger population that the sample was selected from for analysis (Carsey & Harden, 2014). It is very rare to have a large number of different samples from a population to repeat the analysis; hence Monte Carlo simulation should be utilized to create many sampling data with similar parameters of a population.

Carsey and Harden (2014) define Monte Carlo simulation as “any computational algorithm that randomly generates multiple samples of data from a defined population based on an assumed Data Generating Process (DGP)” (p.4). Monte Carlo simulation allows researchers to control the parameters of the population DGP, which allows for the comparison of theoretical models and statistical estimators (Carsey & Harden, 2014).

Simulation can provide a tool to examine and comprehend the results of analysis. In addition, simulation can help researchers to control and adjust the magnitude of effects on their own samples. Simulation is also very handy in some cases; such as when there is no analytic solution and unknown statistical properties of an estimator (Carsey & Harden, 2014).

Carsey and Harden (2014) summarize the steps of Monte Carlo simulation as follows:

- Define the specifications of DGP, which are evaluated in the study. The mathematical formulas and procedures must be stated to produce the data for comparison.
- Pattern out the simulated data that are compared with the true DGP.
- Make changes on theoretical DGP and rerun the simulation to examine the different pattern with new simulations.

Number of Replications

The number of replications in MC studies is equivalent to sample size in empirical research. Harwell et al. (1996) recommended a few thousand replications, at minimum, to minimize sampling variance. Another crucial aspect of MC studies is obtaining reliable results. The number of replications should be from 2000 to 5000 to produce reliable estimates for parameters (Hutchinson & Bandalos, 1997).

Variables for Simulation with Factor Retention

There are several options that should be considered for simulated variables in factor analytic research, including factor pattern loadings, number of items per factor, sample size, and factor intercorrelation (Cho, et al., 2009). The most influential among these is the correlation among factors; pattern loading size, and the number of variables per factor. Crawford et al. (2010) manipulated the number of observations (sample size), number of factors, number of items, factor loadings, and factor correlations and found that the difference between precision of PA-PCA and PA-PAF were minor. Moreover, smaller sample size and smaller factor loading led to over-factoring. In addition, correlated factors led to under-factoring error within PA-PCA. When the number of factors is under estimated, rotation results in a distorted solution. When selecting variables, manageability must be taken in to consideration in terms of time and software resources (Hutchinson & Bandalos, 1997). Moreover, Hutchinson and Bandalos (1997) recommend that if simulated variables were categorical, ANOVA would be the most proper design for analyzing results as inferential statistics.

Sample Size for Factor Analytic Retention

Larger sample sizes are preferable in scale development to minimize the random effects of scale variance (Tabachnick & Fidell, 2001). Although Gorsuch (1983) recommended sample sizes of 100 as an absolute minimum, Comrey and Lee (1992) indicate 100 is poor, and suggested 500 as a very good size for a sample. The size of a sample is not an influential variable alone in factor recovery; different combinations of variables result in a good factor recovery. For example a sample size of 40 would be good enough in some combinations of conditions while a sample of 1000 could be inadequate in certain situations such as low communalities, and large number of weakly determined factors (MacCallum et al., 2001). MacCallum et al., (1999) reported in their sample size study that with high communalities, sample size and over-determination; it is possible that even a very small sample size (n<100) could be sufficient.

The increase in number of factor, ratio of number of variables and sample size leads to better accuracy in estimation of factor loading and better stability in the solution (Browne, 1968). In addition, Zwick and Velicer (1986) advised that the sample size should be at least 5 times the number of variables. The cost of collection and availability of subjects limits the size of the sample. In other words, to estimate population parameters with accuracy, the sample size should be sufficient which is often a limitation imposed by cost and availability.

Polychoric Correlation in Factor Analytic Retention

Even though many educational and psychological measures are composed of Likert scale items, popular statistical packages are performing PCA and factor analysis with Pearson correlation, which means that nominal and ordinal data were treated as interval or ratio data. Polychoric correlation coefficients are maximum likelihood estimates of Pearson correlation with normally distributed variables. It is assumed that the item responses are normally distributed. Even though a Pearson correlation analysis is based on the assumption of a linear relationship between variables, this may not be the case for ordinal data. Due to strong skewness and kurtosis in Likert scale items, it is suggested that using polychoric correlation would be appropriate with ordinal data instead of Pearson correlations (Basto & Pereira, 2012; Bernstein & Teng, 1989).

Timmerman and Lorenzo-Seva (2011) defined polychoric correlation as a maximum likelihood estimate for Pearson correlation between underlying variables. It is assumed that ordered polytomous variables represent an underlying bivariate normally distributed variable. Garrido et al., (2011) indicated that polychoric correlation is more accurate than Pearson correlation in determining the number of factors to retain with ordinal items when using the MAP method. If Pearson correlation were used to calculate correlation coefficients with ordinal items, the strength of the relationship between variables would be underestimated (Cho et al., 2009). Polychoric correlation, which produces unbiased estimates of relationships among variables, is recommended with ordinal data (Cho et al., 2009).

Even though polychoric correlation produces unbiased parameter estimates for EFA, the procedure generates large sampling errors and often produces indefinite correlation matrices (Garrido et al., 2011).

Limitations of Monte Carlo Simulation

While MC simulation provides many advantages, the procedure has disadvantages. First, some permutations of variable combinations do not reflect real life conditions and often are not helpful. Second, improper model design can lead to poor external validity. Third, extreme conditions of simulation studies are similar to outliers in general inferential analysis. Fourth, inadequate numbers of replications affects the dependability of results. Finally, it is often impossible to include all factors as simulated variables that could have an effect on the results (Hutchinson & Bandalos, 1997).

Limitations of MC simulation can be summarized in two points. First, the practicality of MC simulation heavily depends on the reflection of the reality of conditions that are modeled and generated. Second, the number of replications in MC simulation is crucial, and it is important to produce a model that fits the sampling distribution. In other words, because Monte Carlo simulation generates random data with constraints, it is possible to not represent whole points and characters of real population data (Hutchinson & Bandalos, 1997). Additionally, more conditions or parameters of variables can lead to longer simulation runs. It is also difficult to distinguish significant combinations of conditions. One of the criticized aspects of MC studies is that the researcher has full control of the parameters, which could lead to very rare and extreme results in regular research settings (Hutchinson & Bandalos, 1997).

Implications for Education: This study recommends comparing the methods and procedures for determining the number of factors to retain with categorical data, utilizing polychoric correlations via Monte Carlo simulation. For instance, this work might help more teachers, administrators, and students better measure academic performance and reduce the failure rate. These methods will also provide a clearer picture for teachers and administrators to assess their achievements and measure teaching styles. These methods can be implemented in the classroom environment to address various issues related to learning styles, educational performance, and more.

Conclusion

Even though vast amount of measurement instruments in education and psychology consist of items measured with Likert scales generating ordinal data, when these ordinal data are analyzed in factor analysis with Pearson correlation the data are mistakenly interpreted as interval or ratio scale data. Parallel to this general misconception, ordinal and categorical data in factor analysis with polychoric correlation are understudied.

The performance of each retention method should be examined on simulated categorical data. The future studies should also test these procedures to determine relative effectiveness for categorical data under a variety of practical conditions by varying the number of respondents, number of factors in population, magnitude of inter-factor correlations, and factor loadings.

References

Basto, M., & Pereira, J. M. (2012). An SPSS R-menu for ordinal factor analysis. Journal of Statistical Software, 46(4), 1-29.

Beauducel, A. (2001). Problems with parallel analysis in data sets with oblique simple structure. Methods of Psychological Research Online, 6(2), 141–157.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. Psychological Bulletin, 105(3), 467-477.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York, NY: Guilford Press.

Browne, M. W. (1968). A comparison of factor analytic techniques. Psychometrika, 33(3), 267-334.

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. Multivariate Behavioral Research, 27(4), 509–540.

Carsey, T. M., & Harden, J. J. (2014). Monte Carlo Simulation and Resampling Methods for Social Science. Thousand Oaks, CA: Sage.

Cho, S. J., Li, F., & Bandalos, D. (2009). Accuracy of the parallel analysis procedure with polychoric correlations. Educational and Psychological Measurement, 69(5), 748-759.

Choi, J., Kim, S., Chen, J., & Dannels, S. (2011). A comparison of maximum likelihood and Bayesian Estimation for polychoric correlation using Monte Carlo simulation. Journal of Educational and Behavioral Statistics, 36(4), 523–549.

Comrey, A. L., & Lee, H. B. (1992). A first course in factor analysis. Hillsdale, NJ: Erlbaum.

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(7). http://pareonline.net/getvn.asp?v=10&n=7

Cramer, D., & Howitt, D. L. (2004). The Sage dictionary of statistics a practical resource for students in the social sciences. Thousand Oaks, CA: Sage.

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(6), 885–901.

Cudeck, R., & MacCallum, R. C. (Eds.). (2007). Factor Analysis at 100 Historical Developments and Future Directions. Mahwah, NJ: Lawrence Erlbaum Associates.

Dinno, A. (2011). Gently clarifying the application of Horn’s parallel analysis to principal component analysis versus factor analysis. Retrieved from http://doyenne.com/Software/files/PA_for_PCA_vs_FA.pdf

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(3), 272–299.

Fava, J. L., & Velicer, W. F. (1996). The effects of underextraction in factor and component analyses. Educational and Psychological Measurement, 56(6), 907–929.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, 7(3), 286-299.

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. Educational and Psychological Measurement, 71(3), 551–570.

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. Psychological methods, 18(4), 454-475.

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(3), 377-393.

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

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. Journal of Personality Assessment, 68(3), 532–560.

Gorsuch, R. L. (2003). Factor analysis. In J. A. Schinka & W. F. Velicer (Eds.). Handbook of psychology: Research methods in psychology (Vol. 2, pp. 143–164). John Wiley and Sons Inc.

Harwell, M., Stone, C. A., Hsu, T. C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. Applied Psychological Measurement, 20(2), 101–125.

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(2), 191–205.

Henson, R. K., Capraro, R. M., & Capraro, M. M. (2004). Reporting practices and use of exploratory factor analyses in educational research journals: Errors and explanation. Research in the Schools, 11(2), 61–72.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. Educational and Psychological Measurement, 66(3), 393–416.

Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Quality & Quantity, 44(1), 153-166.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30(2), 179-185.

Hutchinson, S. R., & Bandalos, D. L. (1997). A guide to Monte Carlo simulation research for applied researchers. Journal of Vocational Education Research, 22(4), 233–245.

Lorenzo-Seva, U., & Ferrando, P. J. (2006). Factor: A computer program to fit the exploratory factor analysis model. Behavior Research Methods, Instruments, & Computers, 38(1), 88–91.

Lorenzo-Seva ,U.,Timmerman,M.E., & Kiers,H.A.L. (2011). The hull method for selecting the number of common factors. Multivariate Behavioral Research, 46(2), 340–364.

MacCallum, R.C.,Widaman, K.F., Preacher, K.J., Hong, S. (2001). Sample size in factor analysis: the role of model error. Multivariate Behavioral Research. 36(4), 611–637.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. Psychological methods, 4(1), 84-99.

Mooney, C. Z. (Ed.). (1997). Monte Carlo simulation. SAGE.

Mundform, D. J., Schaffer, J., Kim, M. J., Shaw, D., Thongteeraparp, A., & Supawan, P. (2011). Number of replications required in Monte Carlo simulation studies: A synthesis of four studies. Journal of Modern Applied Statistical Methods, 10(1), 19-28.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods, Instruments, & Computers, 32(3), 396–402.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. Structural Equation Modeling, 8(2), 287-312.

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(1), 28–56.

Revelle, W. (2013). psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston. R package version 1.3.9

Revelle, W. & Rocklin, T. (1979). Very Simple Structure - alternative procedure for

estimating the optimal number of interpretable factors. Multivariate Behavioral Research, 14(4), 403-414.

Spearman, C. (1904).“General Intelligence,” Objectively Determined and Measured. The American Journal of Psychology, 15(2), 201-292.

Stevens, J. (2002). Applied multivariate statistics for the social sciences. Taylor & Francis US.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics. Allyn and Bacon.

Thoemmes, F., MacKinnon, D. P., & Reiser, M. R. (2010). Power analysis for complex mediational designs using Monte Carlo methods. Structural Equation Modeling, 17(3), 510-534.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. American Psychological Association.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. Psychological Methods, 16(2), 209-220.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41(3), 321-327.

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 (pp. 41-71). Springer US.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. Multivariate behavioral research, 25(1), 1-28.

Wilcox, R. R. (1988). Simulation as a research technique. In J. P. Reeves (Ed.), Educational research, methodology, and measurement: An international handbook (pp. 134-137). Pergamon.

Wilcox, R. R. (1992). Comparing one-step M-estimators of location corresponding to two independent groups. Psychometrika, 57(1), 141–154.

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(4), 354-365.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research a content analysis and recommendations for best practices. The Counseling Psychologist, 34(6), 806–838.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. Psychological bulletin, 99(3), 432-442.