Running Head: NUMBER OF FACTORS

Strategies for Determining the Number of Factors to Retain in Exploratory Factor Analysis

Michael Stellefson and Bruce Hanik

Texas A&M University

Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, February 7, 2008.

Abstract

When conducting an exploratory factor analysis, the decision regarding the number of factors to retain following factor extraction is one that the researcher should consider very carefully, as the decision can have a dramatic effect on results. Although there are numerous strategies that can and should be utilized when making this decision, researchers most often use the eigenvalue-greater-than-one rule or visual scree test. This trend is troubling, given that other, more sophisticated techniques have been shown to give more accurate appraisals of the number of factors. For this reason, understanding how advanced techniques derive the number of retainable factors is important. The present paper reviews various factor retention rules. Data from a questionnaire development study will be analyzed to make the discussion of the various techniques concrete.

KEY WORDS: factor analysis, statistics, eigenvalues

When conducting an exploratory factor analysis, behavioral scientists often want to represent a large set of measured variables using a smaller, more parsimonious set of latent variables, while still preserving the essential original information contained within the original data (Zwick & Velicer, 1982). Researchers using exploratory factor analysis (EFA) must make a number of analytical decisions to achieve this summary, including selection of: 1) extraction method, 2) number of retainable factors, 3) rotation method following extraction, and in some cases 4) calculation method of factor scores. The decision regarding the number of factors to retain following factor extraction (i.e., #2) is one the applied education researcher should consider *very* carefully (Fava & Velicer, 1992; Hayton, Allen, & Scarpello, 2004), as the decision can have a direct effect on results.

Why is this decision such an important one? The ramifications of under- or over-extraction can dramatically distort subsequent results (Comrey & Lee, 1992; Kaiser, 1960; Zwick & Velicer, 1986). Over-extraction followed by rotation causes minor factors to be built up at the expense of major factors (Zwick & Velicer, 1986). When this occurs, factors normally only possess one large pattern coefficient and a few low pattern coefficients. Factors of this ilk are hard to interpret and unlikely to be replicable (Zwick & Velicer, 1986). Under-extraction followed by rotation can result in a loss of information, because the researcher is either erroneously ignoring one factor or combining one distinct factor with others. While under-extraction is generally considered a more severe error (Fava & Velicer, 1992), the misspecification of factor structure in either way (i.e., retaining too many or too few factors), can increase error (Henson &

Roberts, 2006) and lead to poor factor interpretation and reproduction (Velicer, Eaton, & Fava, 2000).

Incidentally, the decision regarding the number of factors to retain would not be as critical if it were appropriate to accept an unrotated factor solution. If an unrotated factor solution was acceptable, then the structure of all identified factors would not be affected by the number of factors retained, because factors when they are first extracted are always perfectly uncorrelated (Zwick & Velicer, 1982). This truism is scarcely considered, however, as factor rotation is usually *essential* to the interpretation of factor structure (Thompson, 2004).

Qualities of Retained Factors

Identifiable factors are comprised of at least two measured variables with non-zero "pattern/structure" coefficients (Thompson, 2004). Before and after varimax (i.e., orthogonal) rotation, pattern and structure coefficients of factors are referred to interchangeably due to their equivalence in orthogonal rotation (Anderson & Rubin, 1956; Morrison, 1976; Thompson, 2004). Within factor analysis, pattern/structure coefficients are analogous to regression beta weights assigned to predictor variables that are perfectly uncorrelated with one another. Factor pattern/structure coefficients attempt to reproduce relationships among measured variables accounted for within each factor. In principal components analysis, the pattern/structure coefficients of each variable on a given factor are squared and added together to compute that factor's eigenvalue. An eigenvalue is the index of the amount of information represented in a factor (Thompson, 2004).

The larger a factor's eigenvalue, the more variance that it accounts for within a group of measured variables. A factor with an eigenvalue of 1.0 accounts for as much variance as a single variable being perfectly correlated with the factor, while a factor with an eigenvalue over 1.0 provides more summarizing power than one single variable alone. Factors with near 0 eigenvalues provide no summarizing power (Zwick & Velicer, 1986), and thus cannot adequately reproduce any of the relationships among measured variables. Based on these factor analytic principles, the researcher seeks to retain the factors which reproduce important relationships and disregard factors which do not reproduce important relationships among the measured variables of interest.

Factor Retention

There are numerous strategies that can and should be utilized when determining the correct number of factors to retain. Many researchers (Henson & Roberts, 2006), however, automatically pick the eigenvalue-greater-than-one rule (Guttman, 1954) or the visual scree test (Cattell, 1966) over more accurate methods which are available (Fabrigar et al., 1999; Ford, MacCallum, & Tate, 1986; Hayton et al., 2004). This is because many applied researchers automatically accept the "default" analytic choices built into software such as SPSS and SAS. This is also because not all techniques are easily automated via "point and click" procedures available within modern statistical software packages (O'Connor, 2000; Thompson & Daniel, 1996). The reliance on sub-par methods, such as the eigenvalue-greater-than-one rule or visual scree test, is troubling (as we will see), given that other techniques provide more accurate appraisals regarding the ideal number of retainable factors (Thompson & Daniel, 1996; Zwick & Velicer, 1986). Further, Monte Carlo simulations illustrate the superiority of these other, more advanced

strategies, which provide more valid and reliable factors (Zwick & Velicer, 1986). Therefore, it is important to 1) understand how different techniques derive the number of retainable factors and 2) make use of alternative and more sophisticated factor retention techniques.

To illustrate the utility of various factor retention strategies, the present paper examines and explains various rules that can be used in conjunction with the principal components factor extraction method. The rules covered include: Guttman's (1954) eigenvalue-greater-than-1.0 rule; Cattell's (1966) scree test; the standard error (SE) scree test (Zoski & Jurs, 1996); Velicer's (1976) minimum average partial test; Horn's (1965) parallel analysis; and the non-parametric bootstrap (Diaconis & Efron, 1983). A data set will be analyzed using each strategy in order to demonstrate concretely how to conduct and interpret these analytical options. This data set reflects 12 item scores on a questionnaire developed and pilot tested by Chaney and colleagues (2007), which was the first measure of its kind to evaluate undergraduate health education students' perceptions of distance education quality. The specific items chosen for this analysis assessed both general and course specific opinions related to distance education courses taken at the undergraduate level. Intercorrelations between the 12 items are presented in Table 1. The data collected during this pilot study on items, 1a, 1b, 1c, 1d, 2b, 2c, 2d, 2e, 2f, 2g, 2h, and 2i, were reliable ($\alpha = 0.857$).

Eigenvalue-Greater-Than-1.0 Rule

Guttman's (1954) eigenvalue-greater-than-1.0 rule (sometimes erroneously attributed to Kaiser (1960) and called the "K1" rule) is the most common rule used within statistical software packages to determine the number of retainable factors (Henson &

Roberts, 2006; O'Connor, 2000; Thompson, 2004; Thompson & Daniel, 1996; Zwick & Velicer, 1986). However, its pervasiveness among research articles does not reflect its adequacy as a strategy. As discussed earlier, factors are latent constructs created as aggregates of measured variables and so should consist of more than one measured variable (Thompson, 2004). If a factor has an eigenvalue of 1.0, and, within the factor, only one measured variable has a pattern/structure coefficient of 1.0, then all other measured variables will have pattern/structure coefficients of 0 (Thompson, 2004). The rule posits that noteworthy factors, representing multiple measured variables, should have eigenvalues greater than 1.0, which would be the bare minimal expectation if the factor consisted of a single measured variable, albeit a measured variable perfectly correlated with the factor.

Consider the correlation matrix presented in Table 1, which reflects associations among the survey items listed earlier (Chaney et al., 2007). If the researcher were to utilize the eigenvalue-greater-than-1.0 rule to decide how many factors to retain, the SPSS (SPSS, Inc., 2005) syntax presented in Figure 1 would yield the eigenvalues reported in Figure 2 (outlined in square box). Upon interpreting this output, the researcher would retain components I, II, and III, as each possesses an eigenvalue greater than 1.0. These factors represent more variance in the Table 1 correlation matrix than a single variable within the matrix could possibly represent alone. The pattern/structure coefficients of the items on each factor are also presented in Figure 2 to illustrate the correlations between the measured variables and these latent factors. We can see from this chart how each item contributes to the makeup of each latent factor. While all items have rather strong correlations with factor I, only items 1b and 1c contribute to the make

up of factor II, and only items 2f and 2g contribute to the composite of factor III. This analysis reveals the relative supremacy of factor I as compared to factors II and III in terms of how much variance from the original correlation matrix can be reproduced by each factor.

Given that eigenvalues are sample statistics and have sampling error, judgment must be used when using the eigenvalue-greater-than-1.0 rule to determine the number of factors to retain. Strict adherence to this hard and fast rule is *not recommended*; rather, it is suggested that the informed researcher use theory and past exploratory factor analysis research to direct factor retention decision making, because the method consistently overestimates the number of retainable factors (Gorsuch, 1983; Horn, 1965; Linn, 1968; Zwick & Velicer, 1982). In addition, the method refers to eigenvalues reproduced from matrices that place communality estimates in the diagonal rather than ones; this practice has been shown to be erroneous (Gorsuch, 1980). Moreover, regardless of the rule's omnipresence within statistical software packages as the default option, it is *not* the rule of choice to determine the number of factors to retain (Fabrigar et al., 1999; Glorfield, 1995; O'Connor, 2000; Thompson, 2004; Thompson & Daniel, 1996; Zwick & Velicer, 1986).

Cattell's Scree Test

The scree test is a graphical test used for determining the number of factors to retain (Cattell, 1966). Visually appealing graphs are constructed by plotting eigenvalues along the ordinate (y-axis) and factor numbers along the abscissa (x-axis) (Tanguma, 2000). A mountain- or cliff-like graph is produced, because successively extracted factors have successively smaller eigenvalues. The eigenvalues associated with the

factors included on the "mountainous" part of the graph represent solid, noteworthy factors which should be retained; whereas, trivial factors compose the "scree" (rubble of loose rock not attached to mountains), which should be discarded (Thompson, 2004). Figure 3 presents a scree plot of the factors generated from the Table 1 correlation matrix.

Theoretically, a ruler should be used to draw a line through the "elbow" of the graph and along the scree, as is illustrated in Figure 3. The researcher should retain all factors above the line that is drawn. In the Figure 3 scree plot, the data point for factor I is above this line, which indicates that one factor should be retained. This finding contradicts the result generated when implementing the eigenvalue-greater-than-1.0 rule (i.e., that three factors should be retained).

There are several problems with the scree plot method as well, most notably that factor retention interpretations are susceptible to inter-rater reliability problems resulting from different researchers making different interpretations about the correct number of retainable factors within identical graphs (Crawford & Koopman, 1979; Gorsuch, 1983, pp. 166-168). While the scree plot shown in Figure 3 is quite straightforward in terms of assessing the number of retainable factors, there are instances when the scree plot graph possesses several breaks and more than one line can be drawn along the scree (Zwick & Velicer, 1986). There can also be instances where the graph does not even possess an obvious break in the curve (Zwick & Velicer, 1982). These potential problems make the visual scree test especially susceptible to arbitrary judgments regarding the number of retainable factors.

Standard Error (SE) Scree Test

Because of the subjectivity which may hinder reliable interpretations of visual scree graphs, variations to these plots have been proposed. These variations use regression equations to objectively evaluate where the elbow in the scree plot occurs. The SE scree test (Zoski & Jurs, 1996) has been identified as the most effective regression-based variation of the visual scree (Nasser, Benson, & Wisenbaker, 2002). This test calculates SEs of eigenvalues using a sequence of regression analyses, whereby a decreasing number of eigenvalues are input within consecutive regression analyses. The dependent variables are the eigenvalues' actual numerical values, and the predictor variables are ordinal positions 1 to v (v being the number of variables in the matrix of association) of the eigenvalues as they were extracted. The first SE is calculated based on a regression entering all the v eigenvalues, the second entering v-1 eigenvalues, the third entering v-2 eigenvalues, and so on and so forth (Nasser, Benson, & Wisenbaker, 2002). When each descending eigenvalue is removed from successive regression analyses, the SEs of the R^2s of each regression decrease. Once the SE of R^2 drops below the one divided by the number-of-variables criterion (which is used because error variance tends to be inversely related to sample size) following a regression, there are no eigenvalues with large residuals left in the analysis, so factor retention should cease because there is little information left for inclusion in another factor. Alternatively stated, the number of SEs that exceeds the 1 / v criterion indicates the number of factors to retain.

Figure 4 presents an example of a SE scree test using the eigenvalues computed earlier. The syntax is provided to illustrate how the regressions can be conducted within

SPSS. After each regression, the researcher must delete each preceding factor number and corresponding eigenvalue in order to decrease the number of eigenvalues used in successive regressions. Notice that for our example the first two SEs exceed the criterion estimate of n / 12 = .083; therefore, the SE scree test suggests that two factors should be retained. This result does not corroborate either the eigenvalue-greater-than-1.0 rule or scree test.

Velicer's Minimum Average Partial (MAP) Analysis

Velicer's (1976) MAP analysis is based on matrices of partial correlations (O'Connor, 2000; Zwick & Velicer, 1986). For the purposes of factor analysis, partial correlation is the correlation between two variables with the influence of a latent factor(s) removed (Hinkle, Wiersma, & Jurs, 2003). In the first step of the analysis, the first factor is partialed out of the original matrix of association, and the average squared coefficient in the off-diagonals of the reproduced partial matrix of association is computed (i.e., values above and below the diagonal of the reproduced matrix are squared, added together, and the resulting sum is divided by the total number of entries in the entire reproduced matrix). In the second step of the analysis, the first two factors are partialed out of the original matrix of association, and the average squared partial correlation is again computed (O'Connor, 2000). This process is continued for v -1 steps, v being the number of variables within the matrix of association. These computed values are the basis for MAP analysis. After the appropriate number of steps, the averaged squared partial correlations computed at each step are lined up vertically corresponding to the step number in which each value was computed. The number of retainable factors is

determined by the step number in the analysis which resulted in the lowest average squared partial correlation.

Fortunately, modern computer software and SPSS syntax programs have made this computationally difficult task quite simple. There are two ways of conducting a MAP Analysis using SPSS syntax files written by Brian O'Connor (2000) which can be downloaded at http://people.ok.ubc.ca/brioconn/nfactors/nfactors.html. The researcher can either enter a matrix of association directly into SPSS syntax written by O'Connor (2000), or request that SPSS read a matrix of association that has been saved using a correlation file designation (.cor). Either way, the program will perform all of the necessary calculations, derive eigenvalues for each of the v-l factors, and also indicate the numerical values of the minimum average squared partial correlations according to the step number in which it was computed.

Figure 5 depicts the output from a MAP analysis that was performed on data in Table 1 using the O'Connor (2000) syntax. The smallest average squared partial correlation computed was .025490, and it was derived after the first factor was extracted. Consequently, the MAP analysis indicates that the correct number of factors to retain is one. This result is similar to the result produced by the visual scree test, but contrary to the eigenvalue-greater-than-1.0 rule and the SE scree test. Much like the SE scree test (Zoski & Jurs, 1996), however, MAP analysis provides an unequivocal stopping point for factor retention. Monte Carlo simulations have determined that the MAP analysis is more often accurate than the eigenvalue-greater-than-1.0 rule and visual scree test (Zwick & Velicer, 1986).

Understanding MAP Analysis for Heuristic Value

It should be noted that the average squared partial correlation reaches a minimum when the residual matrix most closely resembles an identity matrix. As more factors are extracted, the residual correlation matrix approaches zeroes (Thompson, 2004). Essentially, factors represent compressed information within a matrix of association that is "sanitized" (i.e., useful information within the matrix is retained and excess "trash" information is discarded as residual). A major component of the "trash" variance is the measurement error variance that occurs because scores are never perfectly reliable (Thompson, 2003). This is an important heuristic concept to understand in order to conceptualize the idea that factors attempt to reproduce original matrices of association. The researcher can request that SPSS extract limited numbers of factors (i.e., override the default SPSS extraction rule of all factors with eigenvalues-greater-than-one) to examine the effects of factor extraction on the residual correlation matrices. To override this SPSS default, the researcher can enter the number of desired factors in parentheses next to the line of syntax reading: /criteria factors.

By extracting the factors one by one, the informed researcher can verify that the number of entries in the residual correlation matrix over |.05| begins to decrease as more and more factors are extracted from the original correlation matrix. Consider the chart presented in Table 2, which lists the number of non-redundant residuals with absolute values greater than 0.05 after one, two, and three factors are partialed out of the Table 1 correlation matrix. Notice that the number on non-redundant residuals greater than |.05| decreases as more factors are extracted. From the first to third factor extraction, the number of non-redundant residuals over |.05| has decreased from 47 to 40 to 35. This

means that there is less and less information to extract from the original correlation matrix as more and more factors are extracted. Heuristically, this also illustrates that the residual matrix does approach an identity matrix (with zeros in the off-diagonals and ones on the diagonal) as more and more factors are extracted. SPSS upon request prints the number of non-redundant residuals with absolute values at the bottom of each residual matrix. This type of example should verify for the researcher that 1) information has been "sanitized" and taken out of the original correlation matrix during each successive factor extraction, and 2) there is less "factorable" information left within the residual matrix because of this "sanitization" and factor extraction.

Parallel Analysis

Parallel analysis (Horn, 1965) is considered one of the most accurate and underutilized methods to determine the number of retainable factors (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Henson & Roberts, 2006; Velicer et al., 2000). To conduct a parallel analysis, the researcher takes the data on the measured variables and creates random score matrices of exactly the same rank using the exact same scores in the raw data set. The random score matrix "parallels" the actual data set with regard to the number of cases and variables (i.e., rank), yet the scores themselves are randomly ordered on each variable (O'Connor, 2000). The reason for creating a randomly-ordered matrix of the same rank is to create eigenvalues that take into account sampling error which influences the set of measured variables. If scores are randomly ordered, then the eigenvalues of the randomly ordered matrix will fluctuate around 1.0 as a function of sampling error. The range of fluctuations will be narrower as the sample size is larger and the number of variables is smaller (Thompson, 2004). The eigenvalues derived from

the randomly-ordered data are then compared to the eigenvalues produced from the original data matrix. Factors are retained whenever the eigenvalues from the original data for a given factor exceed the eigenvalues corresponding to the desired percentile (usually the 95th) of the distribution of random data eigenvalues (Cota, Longman, Holden, Fekken, & Xinaris, 1993; Glorfield, 1995). Meaningful factors extracted from actual data should have larger eigenvalues than the parallel eigenvalues obtained from random data (Montanelli & Humphreys, 1976).

SPSS syntaxes have been developed to conduct parallel analyses (O'Connor, 2000; Thompson & Daniel, 1996). Figure 6 presents an example output from a parallel analysis done with the data used to compute Table 1. The parallel analysis was done using SPSS syntax developed by O'Connor (2000) which can be downloaded at http://people.ok.ubc.ca/brioconn/nfactors/nfactors.html. The researcher must input a) the number of cases within the data set, b) the number of variables within the correlation matrix of interest, c) the number of random data sets to generate and draw eigenvalues from, and d) the percentile from which eigenvalues generated from the random data sets will be used for comparative purposes. Some have noted parallel analysis might be difficult to implement (Velicer et al., 2000), but the utilization of O'Connor's (2000) syntax seems to assuage any difficulty completely.

For this particular analysis, 568 cases were analyzed, and 12 variables were included within the data set. SPSS was asked to generate 100 random data sets to be used to create an eigenvalue distribution for each factor. The eigenvalues (computed from these 100 random data sets) were used for comparison against eigenvalues computed from the original data set. The section of output in Figure 6 labeled "Random Data

Eigenvalues" provides the researcher with the eigenvalues generated from the randomly-generated data sets. The column of interest to the researcher is the one in which eigenvalues are listed under the heading "Prcntyle." These values correspond to the 95th percentile of the eigenvalues for each factor that were obtained from the 100 randomly-generated data sets. Stated a different way, for each factor, SPSS has searched the sampling distribution of eigenvalues using the randomly-ordered data for the eigenvalue that corresponds to the 95th largest eigenvalue in the set of 100.

The researcher must now find the factor eigenvalues in this column which are less than the factor eigenvalues from the original matrix. The random data eigenvalue in this column which corresponds to factor I is less than the eigenvalue originally computed earlier. From this parallel analysis, we can vest confidence that the first factor should be retained, while the second factor may also be useful because its random data eigenvalue is only slightly larger (+.00532) than the actual eigenvalue calculated. Thus, this result corroborates the findings from the visual scree test and MAP analysis, but reveals different results than those generated by the eigenvalue-greater-than-1.0 rule and the SE scree test.

It is important to reiterate that parallel analysis is regarded as one of the best methods for deciding how many factors to extract and retain (Zwick & Velicer, 1986); however, one Monte Carlo simulation noted its tendency to under-factor under certain circumstances (Mumford et al., 2003).

Non-Parametric Bootstrap Method

An even more sophisticated approach to parallel analysis is the non-parametric bootstrap method (Diaconis & Efron, 1983), which uses randomly-selected intact cases of

data reflecting values contained in the original data that are randomly sampled with replacement (Thompson, 2004). The number of cases sampled each time equals the number of cases in the original sample. Descriptive statistics (means, standard deviations, skewness and kurtosis coefficients) of the original data and many randomly-selected cases account for sampling error which can impact eigenvalue computation. The eigenvalues for each factor are computed across the repeated resamples, thus creating empirically estimated sampling distributions of eigenvalues for each factor. From the sampling distributions of the eigenvalues that are created, empirically estimated SEs of the eigenvalues are derived which reveal how replicable the mean eigenvalue of each factor could be across samples.

For comparative purposes across samples, the mean eigenvalues derived from each sample for each factor can only be reported after an orthogonal or varimax rotation is performed on each solution following each resample so the factors across samples are occupying like factor space. Comparing results across samples without this orthogonal rotation would be nonsensical, because distinct factors extracted across samples could be extracted in a differential order and be composed of completely different measured variables (Zientek & Thompson, 2007). Confidence intervals (CIs) can be built around the eigenvalues to determine if they subsume 1.0. In this regard, the non-parametric bootstrap method accounts for the sampling error that affects eigenvalue computation. This sampling error is not considered when using the eigenvalue-greater-than-1.0 rule. Theoretically, this should be the best strategy to use to determine the number of retainable factors (Zientek & Thompson, 2007).

Bootstrap estimation of the sampling distribution of eigenvalues requires use of specialized syntax for SPSS, such as that provided by Zientek & Thompson (2007). The interested reader is directed to the bootstrap EFA SPSS syntax developed by Zientek & Thompson (2007) which can downloaded at: http://www.coe.tamu.edu/~bthompson/datasets.htm. Adaptation of this syntax to an original data set allows the researcher to conduct a bootstrap EFA on their own data. This syntax was used to analyze the selected items from the Chaney et al. (2007) data set. Table 3 presents the sample eigenvalues and mean bootstrap results across 500 resamples from this data. In addition, the Figure 7 depicts the empirically estimated sampling distributions for the eigenvalues computed. Across the 500 resamples, the second eigenvalue ranged from 1.08 to 1.46, while the third eigenvalue ranged from 0.94 to 1.29. It is important to test whether eigenvalues are greater than 1.0, so CIs are created around the eigenvalues (Zientek & Thompson, 2007). From these illustrations, we look to see if each eigenvalue's CI subsumes 1.0, in which case the factor(s) is disregarded. These results suggest that 2 factors be retained, which matches the results garnered from the SE scree test yet differs from results generated by all other reported methods.

Discussion

Table 4 presents the results produced using each method reviewed. Notice the variability in results. However, after examining Table 4, we can be relatively certain that one or two factors can be reasonably retained. Given a) the relatively novel nature of the research endeavor from which the data were acquired (Chaney et al., 2007), b) the fact that under-extraction is generally considered a more severe error (Fava & Velicer, 1992), and c) the non-parametric bootstrap method should be the best strategy to use (Zientek &

Thompson, 2007), the proper factor retention decision might be to keep two factors instead of one. However, in order to make the decision between one and two, it would be important to "consider relevant theory and previous research" (Fabrigar et al., 1999, pp. 281).

Of particular importance is the process that was used to get to the decision point. Despite the varying reliabilities and validities of the strategies that are available (see Mumford et al., 2003; Zwick & Velicer, 1986), it is recommended that several strategies (i.e., there are more than the ones reviewed here, see Mumford et al., 2003) be operationalized before making a factor retention decision in a given study. It is best practice to verify the correct number of retainable factors by implementing and reporting multiple strategies, as the simultaneous use of multiple decision rules is appropriate and often desirable (Henson & Roberts, 2006; Thompson & Daniel, 1996). Others have recommended that only parallel analysis and MAP analysis be implemented, with the scree test used as an additional test to corroborate results of both (Velicer et al., 2000). Regardless of how many and which techniques are used, it is vitally important to not simply rely on the technique that is the default within the statistical software package that you are using.

Based on previous Monte Carlo simulation literature (Mumford et al., 2003; Velicer et al., 2000; Zwick & Velicer, 1986), an ordered list of the efficacy of the reviewed rules is suggested only for use when principal components factor extraction is selected:

1. Parallel analysis
2. MAP analysis

3. SE scree test
4. Visual scree test
5. Eigenvalue-greater-than-1.0 rule

However, note that previous Monte Carlo studies have not investigated the bootstrap, which is the most computer-intensive strategy, and theoretically should be best. The literature would benefit from research which concretely investigates the utility of bootstrap EFA through Monte Carlo simulation. Also of particular note is the relegation of the eigenvalue-greater-than-1.0 rule to the 5th and *last* rank. The eigenvalue greater than 1.0 rule is *not* the recommended factor retention rule, and the method clearly overestimated the number of retainable factors in the illustrative example provided. It is disappointing and surprising that the eigenvalue-greater-than-1.0 rule is the default rule of choice in modern statistical software packages, and thus is the most commonly-used decision rule (Henson & Roberts, 2006). The informed researcher will take the weaknesses of this rule into account when determining which technique to use when conducting an EFA. Through the implementation of more sophisticated and accurate alternatives, researchers will be able to more often summarize data accurately and validate the constructs implicit within measured variables.

References

Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. In Neyman, J. (Ed.), Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. Berkeley and Los Angeles: University of California Press.

Cattell, R. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chaney, B. H., Eddy, J. M., Dorman, S. M., Glessner, L., Green, B. L., & Lara-Alecio, R. (2007). Development of an instrument to assess student opinions of the quality of distance education courses**.** *The American Journal of Distance Education*, *21*, 145–164.

Comrey, A. L. & Lee, H. B. (1992). *A first course in factor analysis* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cota, A. A., Longman, R. S., Holden, R. R., Fekken, G. C., & Xinaris, S. (1993). Interpolating 95th percentile eigenvalues from random data: An empirical example, *Educational and Psychological Measurement, 53*, 585-596.

Crawford, C. B., & Koopman, P. (1979). Note: Inter-rater reliability of scree test and mean square ratio test of number of factors. *Perceptual and Motor Skills, 49*, 223-226.

Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. *Scientific American*, *248*, 116-130.

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.

Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research, 27*, 387-415.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The applications of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology, 39*, 291-314.

Glorfield, 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.

Gorsuch, R. L. (1980). Factor score reliabilities and domain validities. *Educational and Psychological Measurement, 40*, 895-897.

Gorsuch, R. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum Associates.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika, 19*, 149-162.

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.

Hetzel, R. D. (1996). A primer on factor analysis with comments on patterns of practice and reporting. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 175-2006). Greenwich, CT: JAI Press.

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*, 393-416.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences*. Boston: Houghton Mifflin Company.

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

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141-151.

Lautenschlager, G. F., Lance, C. E., & Flaherty, V. L. (1989). Parallel analysis criteria: Revised equations for estimating the latent roots of random data correlation matrices. *Educational and Psychological Measurement, 49*, 339-345.

Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. *Psychometrika, 33*, 37-71.

Longman, R. S., Cota, A. A., Holden, R. R., & Fekken, G. C. (1989). A regression equation for the parallel analysis criterion in principal components analysis: Mean and 95th percentile eigenvalues. *Multivariate Behavioral Research, 24*, 59-69.

Montanelli, R. G., & Humphreys, L. G. (1976). Latent roots of random data correlation matrices with squared multiple correlations on the diagonals: A Monte Carlo study. *Psychometrika, 41*, 341-348

Morrison, D. F. (1976). *Multivariate statistical methods* (2nd ed.) New York: McGraw-Hill.

Mumford, K. R., Ferron, J. M., Hines, C. V., Hogarty, K. Y., & Kromrey, J. D. (2003, April). *Factor retention in exploratory factor analysis: A comparison of alternative methods*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Nasser, F., Benson, J., & Wisenbaker, J. (2002). The performance of regression-based variations of the visual scree for determining the number of common factors. *Educational and Psychological Measurement, 62*, 397-419.

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*, 396–402.

SPSS Inc. (2005). *SPSS Version 13.0*. SPSS Inc., Chicago IL.

Tanguma, J. (2000). Determining the number of factors to retain. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas.

Thompson, B. (1988). Program FACSTRAP: A program that computes bootstrap estimates of factor structure. *Educational and Psychological Measurement, 48*, 681-686.

Thompson, B. (Ed.). (2003). *Score reliability: Contemporary thinking on reliability issues*. Newbury Park, CA: Sage.

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

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach*. New York: The Guilford Press.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: An historical overview and some guidelines. *Educational and Psychological Measurement, 56*, 197-208.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 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 R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy*. Norwell, MA: Kluwer Academic.

Zientek, L., & Thompson, B. (2007). Applying the bootstrap to the multivariate case: Bootstrap component/factor analysis. *Behavior Research Methods, 39*, 318-325.

Zoski, K., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement, 56*(3), 443–451.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17*, 253-269.

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

Table 1

Correlation matrix for 12 items from Chaney et al. (2007) data set

	1A	1B	1C	1D	2B	2C	2D	2E	2F	2G	2H	2I
1A	1.00	.427	.403	.678	.313	.442	.297	.430	.244	.254	.446	.397
1B	.427	1.00	.571	.488	.244	.261	.223	.265	.253	.294	.262	.267
1C	.403	.571	1.00	.462	.256	.272	.221	.326	.275	.255	.211	.275
1D	.678	.488	.462	1.00	.311	.380	.290	.364	.277	.252	.386	.397
2B	.313	.244	.256	.311	1.00	.368	.428	.306	.290	.326	.347	.357
2C	.442	.261	.272	.380	.368	1.00	.300	.500	.273	.272	.426	.361
2D	.297	.223	.221	.290	.428	.300	1.00	.265	.201	.215	.337	.316
2E	.430	.265	.326	.364	.306	.500	.265	1.00	.330	.309	.391	.463
2F	.244	.253	.275	.277	.290	.273	.201	.330	1.00	.498	.226	.290
2G	.254	.294	.255	.252	.326	.272	.215	.309	.498	1.00	.294	.302
2H	.446	.262	.211	.386	.347	.426	.337	.391	.226	.294	1.00	.531
2I	.397	.267	.275	.397	.357	.361	.316	.463	.290	.302	.531	1.00

Note: All correlations are significant at the 0.01 level (2-tailed). n = 568.

Table 2

Number of non-redundant residuals with absolute values greater than 0.05 following successive factor extractions

Factor Extraction Number	Number of non-redundant residuals with absolute values greater than 0.05
1	47
2	40
3	35

Extraction Method: Principal Component Analysis

Table 3

Sample eigenvalues and mean bootstrap results across 500 resamples

Sample Eigenvalue	M(BR)	SE	M(BR)/SE
4.7569	4.7564	0.1976	24.0710
1.2139	1.2655	0.0681	18.5829
1.0996	1.0998	0.0610	18.0342
0.8882	0.8855	0.0549	16.1389
0.7062	0.7256	0.0362	20.0556
0.6384	4.7564	0.0343	138.5641
0.5721	0.5778	0.0315	18.3684
0.5417	0.5227	0.0281	18.5750
0.4790	0.4617	0.0278	16.5934
0.4087	0.4071	0.0227	17.9294
0.3959	0.3637	0.0213	17.0349
0.2994	0.2883	0.0237	12.1595

Note: BR = bootstrap result

Table 4

Results derived using each strategy

Method	Number of Retainable Factors
Eigenvalue-greater-than-1.0 rule	3
Visual scree test	1
Standard error scree test	2
MAP analysis	1
Parallel analysis	1
Non-parametric bootstrap method	2

Figure 1

SPSS syntax for factor analysis of Table 1 matrix

```
FACTOR
  /VARIABLES onea oneb onec oned twob twoc twod twoe twof twog twoh twoi
  /MISSING LISTWISE /ANALYSIS onea oneb onec oned twob twoc twod twoe twof
  twog twoh twoi
  /PRINT INITIAL KMO REPR EXTRACTION FSCORE
  /PLOT EIGEN
  /CRITERIA MINEIGEN(1) ITERATE(25)
  /EXTRACTION PC
  /ROTATION NOROTATE
  /METHOD=CORRELATION .
```

Figure 2

Eigenvalues

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	*% of Variance*	*Cumulative %*	*Total*	*% of Variance*	*Cumulative %*
1	4.757	39.641	39.641	4.757	39.641	39.641
2	1.214	10.115	49.757	1.214	10.115	49.757
3	1.100	9.163	58.920	1.100	9.163	58.920
4	.888	7.402	66.322			
5	.706	5.885	72.207			
6	.638	5.320	77.527			
7	.572	4.767	82.294			
8	.542	4.514	86.808			
9	.479	3.992	90.800			
10	.409	3.406	94.206			
11	.396	3.299	97.505			
12	.299	2.495	100.000			

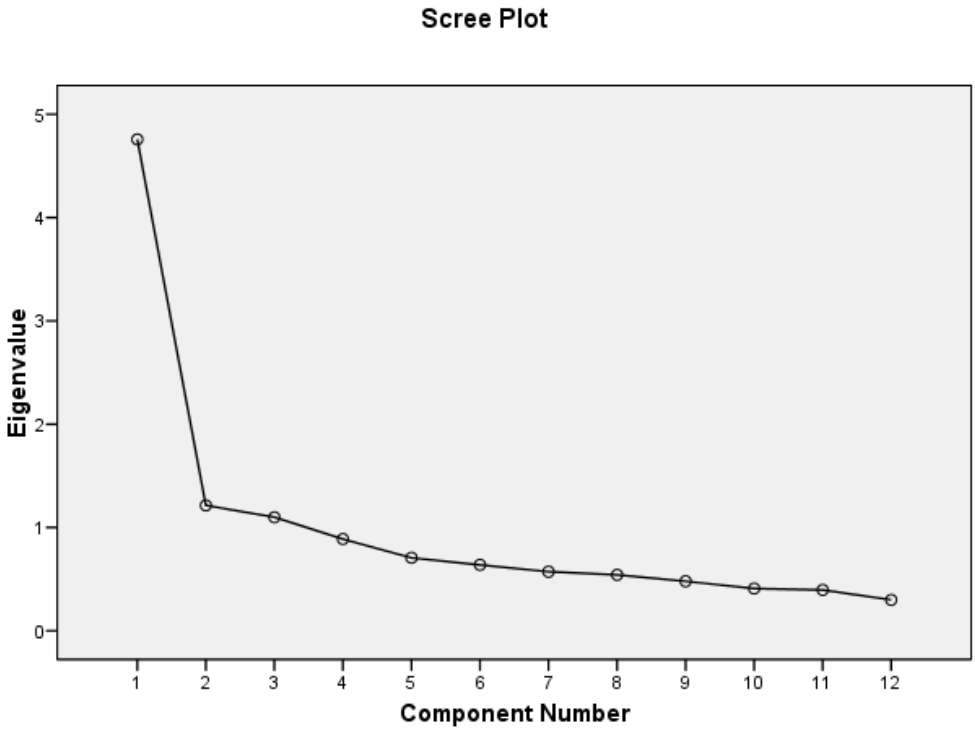
Extraction Method: Principal Component Analysis

Pattern/structure coefficients for each variable on each retained factor

Variable	Component		
	I	II	III
1A	.732	-.238	-.286
1B	.609	-.556	.118
1C	.605	-.532	.139
1D	.723	-.360	-.194
2B	.596	.320	.072
2C	.653	.206	-.170
2D	.527	.269	-.136
2E	.669	.163	-.069
2F	.531	.131	.634
2G	.547	.165	.616
2H	.656	.288	-.285
2I	.666	.253	-.162

Figure 3

Scree plot SPSS output using Table 1 correlation matrix



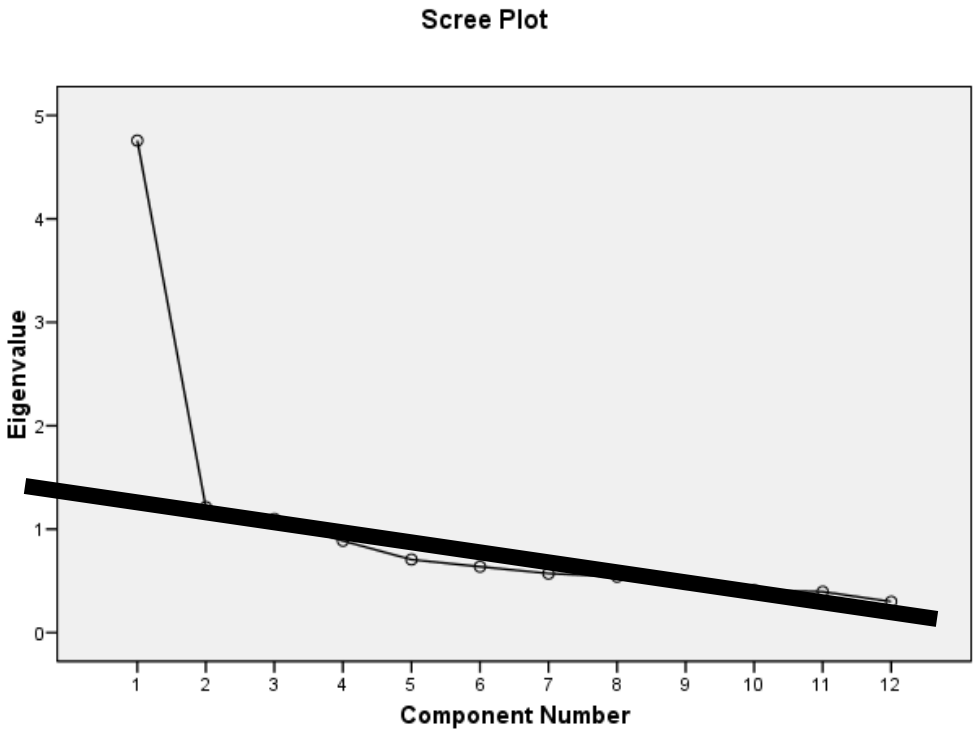


Figure 4

Standard error scree test

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT eigenvalue
  /METHOD=ENTER component  .
```

Number of Factors	R	R Square	Adjusted R Square	**Std. Error of the Estimate**
1	.662	.438	.382	***.95633***
2*	.957	.916	.907	***.09006***
3	.953	.909	.898	***.07786***

Criterion = 1 / 12 = .083

Note: *Number of factors to retain

Figure 5

MAP analysis SPSS output using data on 12 variables

Run MATRIX procedure:

MGET created matrix CR.
The matrix has 12 rows and 12 columns.
The matrix was read from the record(s) of row type CORR.

Velicer's Minimum Average Partial (MAP) Test:

Eigenvalues

```
4.756944
1.213858
1.099601
 .888192
 .706222
 .638383
 .572069
 .541675
 .479001
 .408730
 .395897
 .299429
```

Velicer's Average Squared Correlations

```
     .000000      .123256
    1.000000      .025490
    2.000000      .031490
    3.000000      .040472
    4.000000      .055477
    5.000000      .073832
    6.000000      .094471
    7.000000      .135862
    8.000000      .206703
    9.000000      .381592
   10.000000      .609417
   11.000000     1.000000
```

The smallest average squared *partial* correlation is
.025490

The number of components *to retain* is
1

Figure 6

Parallel analysis SPSS output using data on 12 variables

```
Run MATRIX procedure:

PARALLEL ANALYSIS:

Principal Components

Specifications for this Run:
Ncases    568
Nvars      12
Ndatsets  100
Percent    95

Random Data Eigenvalues

       Root        Means       Prcntyle   Eigenvalues (Derived Earlier)

     1.000000     1.245564     1.300982  <    4.757
     2.000000     1.177730     1.219352  >    1.214
     3.000000     1.127610     1.158623  >    1.100
     4.000000     1.086408     1.117715  >    .888
     5.000000     1.049874     1.077606  >    .706
     6.000000     1.011983     1.038641  >    .638
     7.000000      .977097     1.004292  >    .572
     8.000000      .942732      .970538  >    .542
     9.000000      .909085      .936702  >    .479
    10.000000      .869918      .901895  >    .409
    11.000000      .826360      .857105  >    .396
    12.000000      .775641      .819274  >    .299

------ END MATRIX -----
```

Figure 7

Empirically estimated sampling distribution of eigenvalues of factors from bootstrap factor analysis