

12-1-2009

# Comparison of three computational procedures for solving the number of factors problem in exploratory factor analysis

Adam Vincent Piccone

Follow this and additional works at: <http://digscholarship.unco.edu/dissertations>

---

## Recommended Citation

Piccone, Adam Vincent, "Comparison of three computational procedures for solving the number of factors problem in exploratory factor analysis" (2009). *Dissertations*. Paper 228.

This Text is brought to you for free and open access by the Student Research at Scholarship & Creative Works @ Digital UNC. It has been accepted for inclusion in Dissertations by an authorized administrator of Scholarship & Creative Works @ Digital UNC. For more information, please contact [Jane.Monson@unco.edu](mailto:Jane.Monson@unco.edu).

© 2009

ADAM VINCENT PICCONE

ALL RIGHTS RESERVED

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

A COMPARISON OF THREE COMPUTATIONAL PROCEDURES  
FOR SOLVING THE NUMBER OF FACTORS PROBLEM  
IN EXPLORATORY FACTOR ANALYSIS

A Dissertation Submitted in Partial Fulfillment  
of the Requirements for the Degree of  
Doctor of Philosophy

ADAM VINCENT PICCONE

College of Education and Behavioral Sciences  
School of Educational Research, Leadership, and Technology  
Applied Statistics and Research Methods Program

December, 2009

This Dissertation by: Adam Vincent Piccone

Entitled: *A Comparison Of Three Computational Procedures For Solving The Number Of Factors Problem In Exploratory Factor Analysis*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy in the College of Education and Behavioral Sciences, the School of Educational Research, Leadership, and Technology, Program of Applied Statistics and Research Methods

Accepted by the Doctoral Committee

---

Daniel J. Mundfrom, Ph. D., Chair

---

Jay Schaffer, Ph. D., Committee Member

---

Jamie Perrett, Ph. D., Committee Member

---

Steven Pulos, Ph. D., Faculty Representative

Date of Dissertation Defense \_\_\_\_\_ November 6, 2009

Accepted by the Graduate School

---

Robbyn R. Wacker, Ph. D.  
Assistant Vice President for Research  
Dean of the Graduate School & International Admissions

## ABSTRACT

Piccone, Adam Vincent. A Comparison of Three Computational Procedures for Solving the Number of Factors Problem in Exploratory Factor Analysis. Published Doctor of Philosophy dissertation, University of Northern Colorado, (2009).

Three computational solutions to the number of factors problem were investigated over a wide variety of typical psychometric situations using Monte Carlo simulated population matrices with known characteristics. The standard error scree, the minimum average partials test, and the technique of parallel analysis were evaluated head-to-head for accuracy. The question of using principal components-based eigenvalues versus common factors-based eigenvalues in the analyses was also investigated. As a benchmark, the commonly used eigenvalues-greater-than-one criterion was included. Across all conditions, the principal components-based version of parallel analysis was found to most accurately recover dimensionality using sample correlation matrices drawn from populations with known, simple factor structures. The high degree of accuracy observed for this method suggests that a workable solution to the age-old number of factors problem may be close at hand.

## ACKNOWLEDGEMENTS

“You only get out of an education what you put into it.” This sage advice was offered to me at the pivotal age of 14 by my dad, Peter, when my commitment to institutional learning clearly began to wane. Now, almost 25 years later, I suspect that even he might agree that I have finally put in enough. I would like to thank him and my mother, Francie, for their endless encouragement, and for generously subsidizing my undergraduate degree—a risky leap of faith at the time. I wish to thank Dr. Janice Nerger, Dr. Donald Rojas, and Dr. Lee Rosén, all fantastic instructors at Colorado State University, who were somehow able to envision me as a successful graduate student, and recommended me for graduate school admission.

I am forever indebted to Dr. Paul Retzlaff, not only for serving as advisor and chair for my Master’s Degree thesis project, but for serving as a Doctoral Committee member through comprehensive oral examinations. He is by far the most entertaining instructor I have ever had. His first, droll words to begin every class were, “Thank you all for coming today, I know what a burden it is.” Best wishes for a happy retirement.

Speaking of outstanding teachers, I would be remiss not to mention the exceptional instructors in the Applied Statistics and Research Methods program. Professors like Dr. Susan Hutchinson, Dr. Jamis Perrett (resident SAS expert), and my tireless research adviser, Dr. Daniel Mundfrom, set a lofty standard. Beginning with a persuasive recruitment at the conclusion of one of Dr. Retzlaff’s Psychometrics classes,

Dr. Mundfrom has guided me through every step of this rigorous program. He has been a paragon of erudition, inspiring me to supplement my early program studies with a more mathematically-oriented curriculum. He thoughtfully encouraged me to tackle the more quantitative Applied Statistics concentration, better preparing me for my chosen career.

I also wish to recognize another outstanding educator, Dr. Jay Schaffer, and thank him for his gentle encouragement at numerous critical stages of my graduate education. I would have never even considered taking on the Applied Statistics concentration if Dr. J. was not annually scheduled to teach Mathematical Statistics I and II. Simply put, he is the most well-organized and well-prepared instructor I have ever encountered—all my knowledge of the dark side I owe to him. Thanks also to Dr. Steven Pulos, who so kindly suggested the insightful topic of this dissertation, and who patiently awaited any word as to its progress. Not to be forgotten, I would like to thank my fellow graduate students, especially Emily Pollock and Dr. Robert Pearson. If not for Emily and Bob's help and moral support during our marathon study sessions for comps, none of this could have ever come to fruition.

Finally, I would like to express my undying gratitude to my patient wife, Amanda. With nothing except encouragement and never-ending support, she has been rooting me on as my number-one-fan since the beginning of my protracted graduate tenure. Never pressuring, nor issuing ultimatums to finish up already, sweet Amanda was always in my corner, cheering me on every painful step of the way. Thank you for being such a loyal and steadfast friend. You are the one and only true love of my life.

As far as two certain furry felines go, Max and Daphne, I respectfully thank you both for nothing.

## TABLE OF CONTENTS

### CHAPTER

I.	INTRODUCTION .....	1
	Background .....	1
	Methods for Determining the Number of Factors .....	5
	Delimitation of the Present Study .....	9
	Purpose of the Present Study .....	16
	Research Questions .....	17
	Assumptions and Limitations.....	18
	Significance of the Study.....	19
	Definition of Terms .....	20
II.	REVIEW OF LITERATURE.....	21
	Factor Analysis in the Social Sciences .....	21
	Principal Components versus Common Factor Analysis .....	26
	Importance of Rotating the Correct Number of Factors.....	31
	Methods for Solving the Number of Factors Problem .....	32
	Use of the Number of Factors Methods in the Literature.....	40
III.	METHODOLOGY .....	42
	Monte Carlo Simulation Rationale.....	42
	Independent Variables .....	43
	Other Design Specifications .....	56
	Summary of the Study Design .....	58
	Computer Program .....	59
	Data Generation.....	59
	Data Analysis .....	61
IV.	RESULTS .....	63
	Population Results .....	63
	Sample Results .....	65
	Effects of Specific Design Conditions.....	67



V. DISCUSSION .....	80
Resolution of Research Questions.....	80
Performance of the Procedures .....	84
Study Limitations and Suggestions for Future Research.....	87
Implications and Recommendations.....	89
REFERENCES .....	92
APPENDIX A.....	108

## LIST OF TABLES

Table 1. Varied Saturation Factor Loadings (High, .6 to .8) .....	47
Table 2. Varied Saturation Factor Loadings (Low, .3 to .5).....	48
Table 3. Varied Saturation Factor Loadings (Wide, .3 to .8).....	49
Table 4. Example of a Population Target Matrix with Varied Saturation and No Unique Indicator Variables (Loadings = Wide) .....	51
Table 5. Example of a Population Target Matrix with Varied Saturation and Unique Indicator Variables Present (Loadings = Wide) .....	52
Table 6. Design Conditions for Sample Data .....	55
Table 7. Overall Accuracy of the Number of Factors Methods.....	64
Table 8. Accuracy of PA when the Number of Factors is Large .....	74

## LIST OF FIGURES

Figure 1. Mean hit rates (+ <i>SE</i> ) over all conditions by range of factor loadings (principal components-based methods). .....	68
Figure 2. Mean hit rates (+ <i>SE</i> ) over all conditions by range of factor loadings (common factors-based methods). .....	69
Figure 3. Mean hit rates (+ <i>SE</i> ) across all conditions by number of factors (principal components-based methods). .....	70
Figure 4. Mean hit rates (+ <i>SE</i> ) across all conditions by number of factors (common factors-based methods). .....	71
Figure 5. Mean hit rates (+ <i>SE</i> ) across all conditions by interfactor correlations (principal components-based methods). .....	72
Figure 6. Mean hit rates (+ <i>SE</i> ) across all conditions by interfactor correlations (common factors-based methods). .....	73
Figure 7. Mean hit rates (+ <i>SE</i> ) across all conditions by presence of unique indicator variables (principal components-based methods). .....	75
Figure 8. Mean hit rates (+ <i>SE</i> ) across all conditions by presence of unique indicator variables (common factors-based methods). .....	76

Figure 9. Mean hit rates ( $+SE$ ) across all conditions by number of items per factor (principal components-based methods). .....	77
Figure 10. Mean hit rates ( $+SE$ ) across all conditions by number of items per factor (common factors-based methods). .....	78

## CHAPTER I

### INTRODUCTION

#### Background

Factor analysis, especially the particular incarnation broadly known as exploratory factor analysis (EFA), has long been recognized as an indispensable statistical and psychometric tool (Thompson & Daniel, 1996). Basically, the purpose of this technique is to describe the covariance relationships of a large number of observed random variables in terms of a relatively smaller number of unobserved random quantities, usually called factors<sup>1</sup> (Johnson & Wichern, 2002). Since its modern beginnings in the early 20th century, factor analysis has been closely tied to the field of psychometrics. Karl Pearson, Charles Spearman, and Louis Thurstone pioneered the modern application of this computationally intensive procedure in research focusing on the measurement of abstract psychological concepts such as intelligence. The ability to quantify aspects of intangible constructs proved highly attractive to psychologists. Later, the advent of computer technology and easy-to-use statistical software (e.g., SAS and SPSS) sparked an explosion of factor analytic research in all branches of the social sciences. Perhaps too expansive, in fact, has been the proliferation of factor analytic research to this point, for it has been suggested that researchers have begun to rely more

---

<sup>1</sup> Throughout this dissertation, the term “factor” will be used primarily in its less formal sense, and the more strict distinction between the terms “factor” and “component” may not be implied.

upon default software settings than upon sound statistical methodology (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Preacher & MacCallum, 2003).

One reason for this reliance may be the noticeable lack of a clear consensus as to what constitutes sound factor analytic methodology. Illustrative of this point is the ongoing battle in the literature over the merits of principal components extraction versus extractions based on the common factor model. At its pinnacle, large portions of methodological journals (even entire issues) were devoted solely to the resolution of this feud (e.g., *Sociological Methods and Research*, Volume 17, Issue 4, 1989; *Multivariate Behavioral Research*, Volume 25, Issue 1, 1990). This controversy and others will be expounded upon in greater detail in subsequent sections. Adding to the general confusion among researchers have also been various methodological setbacks, such as the unfortunate rise of overly simplistic heuristics like the “Little Jiffy” approach advocated by Kaiser (1970). Little Jiffy advises retaining all factors with eigenvalues greater than unity, using a principal components extraction algorithm, and employing a varimax rotation prior to interpretation of the solution. Despite shortcuts in the vein of Little Jiffy, the inconvenient fact remains, as Fabrigar et al. (1999) put it, “Perhaps more than any other commonly used statistical method, EFA requires a researcher to make a number of important decisions with respect to how the analysis is performed” (p. 273). Those authors go on to identify no less than five major methodological choices that face the diligent factor analyst, each of which can have major implications for the results of the study. These include choices concerning the size and make-up of the sample, the variable selection stratagem, the model fit algorithm, the method of rotation, and the number of factors to include in the model.

### *The Number of Factors Problem*

Of the multiple decisions confronting the factor analyst, perhaps the most interesting is the question of the number of factors or components to extract prior to rotation. According to Zwick and Velicer (1986), “The determination of the number of components or factors to retain is likely to be the most important decision a researcher will make” (p.432). In the psychometric literature, this decision has become commonly referred to as the number of factors “problem” (Glass & Taylor, 1966, p. 566). The negative connotation of this particular nomenclature is reflective of the harmful impact of extracting an improper number of factors or components. These consequences, which will be outlined in greater detail later, include the appearance of split factors, the emergence of false factors, and the calculation of unreliable factor loadings. Recent extensions of the factor analysis model, such as confirmatory factor analysis, do have some utility when applied to the number of factors question. However, as Hubbard and Allen (1987a) mention, “...it is instructive to note that confirmatory factor analysis has not solved the [number of factors] problem (p. 173-174). Researchers must resist the temptation to interpret the word “confirmatory” in too literal a sense. Instead, it is important to remember that data can never positively confirm a model; they can only fail to disconfirm one (Popper, 1959). So apprehensive are methodologists about widespread over-reliance on the ostensibly named “causal” models, that some have found it necessary to restate such elementary views of scientific methodology explicitly in the context of confirmatory factor analysis (e.g., Cliff, 1983). The results of a properly conducted confirmatory factor analysis, for example, may reveal that a particular three-factor model is a good fit for a set of data. The correct interpretation of such a result is

only that this three-factor model cannot be disconfirmed by the data. However, innumerable other models, including ones with alternative dimensionality, have also failed to have been disconfirmed.

### *Statement of the Problem*

In psychometrics, the number of factors or components extracted in EFA relates expressly to the dimensionality of a psychometric instrument. Although theory and prior research normally dictate the dimensionality of any psychometric instrument, it can be unclear in practice just how many dimensions or factors might actually subsume a particular set of items. This dilemma is especially relevant in the case of test construction and development. Through the years, various recommendations for determining the number of factors to extract have been made. As will be discussed later, most are simply inadequate. Unfortunately, some clearly weak decision rules still pervade mainstream research methodology. For instance, the infamous eigenvalues-greater-than-one rule (Kaiser, 1960), which has been roundly criticized in the psychometric literature for years, still remains the most widely used decision rule in the social sciences (Conway & Huffcutt, 2003; Fabrigar et al. 1999; Ford, MacCallum, & Tait, 1986). The persistent use of Kaiser's rule is almost certainly because it is easy to use, and because it remains the default option in most popular statistical software packages. In fact, it is usually the only option (other than the express specification of a hard number) available to software users—unless custom programming statements are added. Alternative computational methods do exist, however, and are gradually becoming both easier to implement and less time-consuming. The present study will evaluate and directly compare the leading computational methods for determining the number of factors to extract in EFA. This



evaluation and comparison will focus on situations and covariance structures commonly encountered in applied psychometrics.

## Methods for Determining the Number of Factors

### *Statistical Approaches*

Myriad solutions to the number of factors problem have been proposed through the years. Statistical approaches, such as the maximum likelihood ratio test (Lawley, 1940, 1941) and Bartlett's Chi-square significance test (1950, 1951), are among the earliest. However, since these procedures are functions of sample size, even trivial factors can become highly significant when samples are large (Gorsuch, 1973; Harris & Harris, 1971; Hubbard & Allen, 1987a; Jackson & Chan, 1980). The use of these two tests for determining the number of factors to retain is no longer recommended (Velicer, Eaton, & Fava, 2000).

### *Eigenvalues-Greater-than-One Rule*

As previously mentioned, the eigenvalues-greater-than-one rule proposed by Kaiser (1960) has achieved virtual sacrosanctity among researchers despite being repeatedly shown to be unsatisfactory (Browne, 1968; Cattell & Vogelmann, 1977; Hubbard & Allen, 1987a; Gorsuch, 1997; Jackson, 1993; Lee & Comrey, 1979; Linn, 1968; Revelle & Rocklin, 1979; Tucker, Koopman, & Linn, 1969; Yeomans & Golder, 1982; Zwick & Velicer, 1982, 1986). Kaiser first suggested this rule in a famous speech to APA members in September of 1959, and presented it as "...a 'best' answer to the question of the number of factors" (Kaiser, 1960, p.145). He based this idea mainly upon his own personal experience, but also upon a generalization of a well-known algebraic result by Guttman (1954). Since that time, the rule has become widely known as the

Kaiser-Guttman criterion, although it is now clear that Kaiser took Guttman's proof (which considered only population correlation matrices) too far when he extended it to include sample correlation matrices (see Schönemann, 1990). As Cliff (1988) explains, "Sampling effects tend to increase the number of eigenvalues greater than one, so the rule tends to overestimate the number of factors when applied to sample matrices" (p. 279). Further erroneous support for the eigenvalue-one rule stemmed from Kaiser's contention that for a component to have positive KR-20 reliability (Kuder & Richardson, 1937), its eigenvalue must exceed one. However, Cliff argues, "...Kaiser's rationale for relating the reliability of components to the number of eigenvalues greater than unity was based on a misapplication of a common formula for the reliability of a composite..." (p. 276). Cliff goes on to demonstrate the reliability of a principal component depends upon the reliability of the measures, not upon the size of eigenvalues. Through time, then, the theoretical rationale behind Kaiser's recommendation has severely eroded. Meanwhile, empirical evidence exposing the weaknesses of this rule has mounted. In fact, in the wake of a virtual maelstrom of published denunciation, Preacher and MacCallum (2003) were compelled to remind readers, "The Kaiser criterion will at least occasionally yield a correct estimate of the number of factors to retain" (p. 29). Currently, though, most authors (including Preacher and MacCallum) call for the abandonment of the eigenvalues-greater-than-one criterion for determining the number of factors (Hubbard & Allen, 1987a; Merenda, 1997; Velicer et al., 2000).

#### *Parallel Analysis*

Horn (1965) was one of the first to recognize that Kaiser's (1960) eigenvalues-greater-than-one criterion would not properly account for the effects of sampling. Horn

demonstrated that, on average, the latent roots extracted from normally distributed random data will be greater than unity for half of the total number of roots, simply due to the effects of sampling. This fact alone seriously undermines the theoretical foundation of the Kaiser criterion. To account for sampling error, Horn proposed the technique of parallel analysis (PA). In PA, the researcher compares eigenvalues calculated from the sample data to eigenvalues calculated from one or more sets of random data with the same number of variables and observations. Instead of using unity as the cutoff criterion, only those factors with observed eigenvalues greater than the corresponding eigenvalues from random data are retained. Unfortunately, Horn's sample-based version of the Kaiser rule proved to be too far ahead of its time; the computing power necessary to conduct PA properly would not be broadly available for decades. Even now, implementing PA is moderately difficult programmatically. So, Horn's idea languished as easier solutions to the number of factors problem flourished. Nonetheless, PA has recently resurfaced, enjoying growing recognition as one of the best approaches to the number of factors problem (Conway & Huffcutt, 2003; Gorsuch, 2003; Hayton, Allen, & Scarpello, 2004; Lance, Butts, & Michels, 2006; Merenda, 1997; Thompson & Daniel, 1996; Velicer et al., 2000).

### *Scree Test*

Also unhappy with Kaiser's eigenvalues-greater-than-one criterion, yet hoping to find a method with similar ease of use, Cattell (1966) developed the visual scree test. Conceptually, the visual scree can be thought of as a geometric analogue to Bartlett's Chi-square test (Horn & Engstrom, 1979). This popular procedure is indeed easy to use, but has received mixed criticism in the literature (Cattell & Vogelmann, 1977; Cliff,

1970; Crawford & Koopman, 1979; Jackson, 1993; Kanyongo, 2005; Streiner, 1998; Tucker et al., 1969; Zwick & Velicer, 1982). In a comprehensive review of several number of factors procedures, Velicer et al. (2000) recommend the use of the visual scree only as an adjunct procedure. Even though great pains have been taken to outline a specific, standardized approach to the visual scree (Cattell, 1978; Cattell & Vogelmann, 1977), researchers still must choose a cut-point visually. This inherent subjectivity has been seen as a major weakness (Finch & West, 1997; Humphreys & Montanelli, 1975; Streiner, 1998). For example, Finch and West complain, “The primary problem with the scree test is that it is an ‘eyeball test’; the point of the break in the plot can be difficult to determine or there may be more than one such break” (p. 466). Over the years, several objective methods for determining the cut-point utilizing regression-based approaches have been proposed (Gorsuch & Nelson, 1981; Zoski & Jurs, 1993, 1996). Since these new methods remove the subjectivity component, and since they can be easily automated for computer applications, they are quite promising in their utility. Recently, Nasser, Benson, and Wisenbaker (2002) evaluated the performance of four of these regression-based variations of the visual scree test in terms of ability to identify the true number of factors in sample correlation matrices drawn from populations with a known structure. The standard error scree (SEscree) procedure of Zoski and Jurs (1996) emerged as the most accurate of the regression-based methods, and performed comparably to the traditional visual method. Although use of this promising new procedure is still rare in practice (e.g., Natsopoulos, Kiosseoglou, Xeromeritou, & Alevriadou, 1998) methodologists in the social sciences have started recommending use of the SEscree to

help solve the number of factors problem (Benson & Nasser, 1998; Goodwin & Goodwin, 1999).

### *Minimum Average Partial Test*

Another potential solution to the number of factors problem is the minimum average partial (MAP) test suggested by Velicer (1976). In MAP, the researcher makes use of the off-diagonal elements of partial correlation matrices. The partial correlation matrix is the matrix of correlations that remain after successive factors are partialled out of the sample correlation matrix. In its original form, the average of the squared partial correlations is taken for every principal component. The component number corresponding to the minimum of these averages indicates the number of non-trivial components that should be extracted. MAP can be conceptualized as a measurement of the size of residuals left over after successive numbers of factors have been removed. In this way it is similar to the model fit indices used in CFA, such as those described by Bentler (1989). In fact, Gorsuch (2003) calls MAP “an index of the residuals in EFA” (p. 159). Velicer simply standardizes the residuals by converting them to partial correlations. In simulation studies, MAP performs superbly, second only to PA in accuracy (Peres-Neto, Jackson, & Somers, 2005; Zwick & Velicer, 1982, 1986).

### *Delimitation of the Present Study*

#### *Candidates for Comparison*

Of the countless heuristics for determining the number of factors to extract in EFA, only a few have survived the rigors of scientific scrutiny and emerged as the most robust. To date, the three most reliable methods for determining the number of factors in exploratory factor analysis are the PA technique developed by Horn (1965), the MAP test

suggested by Velicer (1976), and the SEscree test proposed by Zoski and Jurs (1996). The first two methods are older, and have thus faced the most scrutiny; but, they have consistently emerged as accurate (Velicer et al., 2000; Zwick & Velicer, 1986)—usually much more accurate than other methods. The third method, SEscree, is newer. But, it is the most promising of the automated versions of Cattell's (1966) popular visual scree.

### *Direct Comparison of PA, MAP, and SEscree*

As mentioned above, regression-based analogs of the visual scree test are emerging as viable solutions to the number of factors problem in EFA. Nasser et al. (2002) compared the performance of some of these regression-based variations in a Monte Carlo simulation study. In the past, simulation studies have been extensively used to compare the performance of different stopping rules (e.g., Hubbard & Allen, 1987a; Velicer, Peacock & Jackson, 1982; Zwick & Velicer, 1982, 1986). These types of simulations typically involve the generation of a population correlation matrix from a pattern matrix with a known common factor structure. Then, sample correlation matrices can be drawn from the population correlation matrix, usually by means of the Kaiser-Dickman (1962) method. The Nasser et al. study used methodology similar to previous research, and concluded that the SEscree was the most accurate of the regression-based methods. Furthermore, the SEscree test was determined to have accuracy comparable to the traditional visual method. These results suggest that the SEscree may be on par with proven procedures such as PA and MAP. Unfortunately, as Nasser et al. ultimately concede, a major limitation of their work was that these three methods were never directly compared. One purpose of the present study is to extend the results of Nasser et al. by directly comparing the SEscree procedure to PA and MAP in a similar Monte Carlo

simulation. Additionally, the results of the commonly used eigenvalues-greater-than-one rule were calculated for comparison.

### *Population Correlation Structure*

Another limitation of Nasser et al. (2002) involved the structure of population matrices from which the Monte Carlo samples were drawn. Velicer et al. (1982) define a useful nomenclature for these population “target patterns.” In their terms, only “ideal” patterns in the population matrix were examined by Nasser et al. That is, all variable loadings were exactly equal in size, the patterns were “simple” in structure, and all variables loaded on exactly one factor. Of course, this circumstance would probably never present itself in practice. A more realistic situation to expect would be for the loadings to vary in saturation from large to small, and for “unique” indicator variables to sometimes exist in the data. So, in the interest of practicality, only varied saturation population patterns were examined in this study.

Thurstone (1935) introduced the concept of simple structure as it relates to factor analysis. Basically, this is the idea that all indicator variables comprising a psychometric instrument should have exactly one salient factor loading on exactly one of the factors underlying the instrument. As a consequence, the measure of each factor is comprised strictly of items that have salient loadings on the underlying factor, and trivial loadings on all other factors. By this definition, any items that do not have a salient loading on any factor, or have non-trivial loadings on more than one factor are poor items, and should be rewritten, rethought, or discarded altogether during test development. As Stewart (1981) puts it, “The ultimate goal of any factor analysis should be the identification of not only interpretable factors but also simple structure” (p. 61).

Often, in practice, researchers will discover some indicator variables (which are expected to load on one and only one factor) may actually have salient loadings on multiple factors, or perhaps on none. In their work comparing principal components extraction to common factor analysis, Velicer et al. (1982) consider what they call “complex” population target matrices—that is, ones which contain indicators with multiple factor loadings. They also include items which do not load on any factor in the population, the so-called unique variables. Wood, Tataryn, and Gorsuch (1996) found that the addition of unique variables was important when examining the effects of over- and under-extraction in factor analysis. In this study, then, the effect of including unique variables was investigated as it relates to the accuracy of the methods under examination. Complex population targets; however, were not included in the present design. Instead, population targets with simple structure remain the focus of this dissertation.

### *Principal Components versus Common*

#### *Factor Analysis*

No other issue related to factor analysis elicits more passionate controversy than the question of whether to employ a principal components or a common factor model. This subject will be discussed in greater detail later, but it brings up a related methodological question. Almost all of the published controversy surrounding this issue focuses on which model should be used for the actual analysis (i.e., after the number of factors to extract has been determined). Relatively little has been written about which model should be used for deciding the number of factors or components to retain. Generally, methods for determining the number of factors to extract use eigenvalues derived from the sample correlation matrix. In principal components analysis, this



correlation matrix always contains unities in the diagonals. Guttman (1954) refers to this matrix as the unreduced correlation matrix. By definition, factor analytic procedures utilize a correlation matrix with communality estimates inserted into the diagonals with values less than unity. Guttman calls this matrix the reduced correlation matrix.

Since little has been written concerning the choice of matrix from which to compute the eigenvalues for computational methods of determining the number of factors, a common assumption may be that one should just use the same model during both phases of the analysis. Some (e.g., Hakstian & Muller, 1973; Hubbard & Allen, 1987a; Humphreys & Illgen, 1969; Montanelli & Humphreys, 1976; Widaman, 1993) have argued that if the eventual goal is to conduct a common factor analysis, then communalities should be placed on the diagonal of the correlation matrix used to produce the eigenvalues examined to determine the number of factors. However, some rules, like the famous eigenvalues-greater-than-one criterion and the visual scree test were originally developed using eigenvalues from the unreduced correlation matrix. Additionally, many credible methodologists have offered advice specifically contrary to using only the reduced correlation matrix to solve the number of factors problem before conducting a common factor analysis.

A notable example of this advice comes from Johnson and Wichern (2002). In a section entitled “A Strategy for Factor Analysis” these authors are quick to mention the choice of the number of factors as the single most important decision facing the exploratory factor analyst, yet offer no methodology for arriving at even a practical starting point. They do, however, advise performing “a principal component factor analysis” on the data as the first step in their strategy. They explain, “This method is

particularly appropriate for a first pass through the data” (p. 517). As rationale for their claim, Johnson and Wichern make the point that for principal component analysis, it is not required that the sample correlation matrix be nonsingular. A method based on the principal components from a first pass could be performed concurrently to estimate the number of factors, giving at least a reasonable starting point for subsequent analyses. Cliff (1988) suggests just such an approach, first using a method based on principal components as a basis for solving the number of factors problem, and then using a common factor model for subsequent analysis and the estimation of loadings. This methodology is also advocated by Velicer et al. (2000).

Interestingly, a careful examination of the seminal articles describing the visual scree also lends credence to this strategy. Cattell and Vogelmann (1977) introduce the visual scree test thusly:

To be precise about the scree procedure, it consists in entering unities in the diagonal of the given correlation matrix and extracting successive latent roots by a principal axis program down to  $n$  roots, when  $n$  is the number of variables (the last root may be zero). (p. 292).

So, the original scree test is a procedure, based on the eigenvalues from an unreduced correlation matrix, used to provide insight on the number of factors problem prior to a common factor analysis. However, a researcher might choose to construct the scree plot using eigenvalues computed from a reduced correlation matrix. In fact, the SEscree procedure has been applied in exactly that way (Nasser, 1997; Nasser et al., 2002). Interestingly, both PA and SEscree can be based on eigenvalues from either the unreduced or a reduced correlation matrix. Similarly, Gorsuch (1990) claims that the MAP test—which was most decidedly developed using eigenvalues from the unreduced correlation matrix—can be adapted into a common-factor based procedure (p. 33). He

suggests simply computing MAP using residuals from a common factor analysis of the reduced correlation matrix (R. L. Gorsuch, personal communication, February 18, 2008).

Thus, all three procedures under examination here can be based on either a principal components or a common factor model; but, which model should the researcher choose when determining the number of factors to retain during EFA? A thorough review of the research literature failed to reveal a single previous empirical investigation into this question. Here, it will be argued that the model which provides the most accurate solution to the number of factors problem should be used as a first step, regardless of which model is ultimately used for the actual analysis. To that end, the effects of applying the common factor model versus using principal components when determining the number of factors to retain will be compared in the present study by varying which correlation matrix, a reduced or an unreduced one, is used in the model.

As mentioned previously, the eigenvalues-greater-than-one rule is to be used as a benchmark for comparison in the current examination. As such, it will be calculated for every population and sample correlation matrix generated in the study. It is important to reiterate here that this rule, unlike the other three methods, is clearly misapplied when eigenvalues from the reduced correlation matrix are used (see Fabrigar et al., 1999, p. 278; see also Gorsuch, 1980; Gorsuch, 2003; Guttman, 1954; Horn, 1969). However, for comparison and in the interest of completeness, the eigenvalues-greater-than-one criterion was applied to eigenvalues computed from both the unreduced (principal components) and the reduced (common factor) correlation matrices.

### Purpose of the Present Study

The purpose of the present study was to directly compare the leading computational methods for determining the dimensionality of a population covariance matrix based on a sample covariance matrix. This evaluation was based on each method's ability to indicate the correct number of factors in population covariance matrices with known factor structures, using samples generated through Monte Carlo simulation. The effect of adding unique indicators was considered, and the magnitude of item loadings was systematically varied. Conditions that closely resemble a wide variety of psychometric situations were simulated. The present examination also provides the most comprehensive empirical evidence to date regarding the use of unreduced versus reduced correlation matrices, and its effects on the accuracy of the methods.

## Research Questions

In an effort to succinctly outline the goals of the present work, six specific research questions were carefully constructed. This dissertation systematically investigates the following:

- Q1 Do any of the three leading computational procedures have difficulty identifying the dimensionality of population correlation matrices with known, simple factor structures?
- Q2 Which of the three procedures most accurately recovers dimensionality using sample correlation matrices drawn from populations with known, simple factor structures?
- Q3 How does the eigenvalues-greater-than-one rule perform in comparison to the other procedures?
- Q4 How does the addition of unique (noise) indicator variables affect the performance of the methods?
- Q5 Does any single procedure perform well over a wide variety of typical psychometric situations?
- Q6 What effect does using the reduced correlation matrix (i.e., applying a common factor model) versus using the unreduced correlation matrix (principal components) have on the accuracy of the methods?

### Assumptions and Limitations

First, this dissertation focuses only on combinations of variables, loadings, and interfactor correlations that closely resemble conditions commonly encountered in applied psychometric settings. These conditions may not generalize to some other applications of EFA in other disciplines, such as genetic research. Second, the present design focuses specifically on population target patterns with simple structure. In practice, items will often have more complex loading patterns, with salient loadings on multiple factors. Next, although countless methods for estimating communalities under the common factor model exist, the method of principal axis common factoring was selected to represent common factor analysis. As such, the reduced correlation matrix associated with the use of this procedure was the basis of all calculations of common factors-based eigenvalues. Additionally, the current study employs only variables that, when standardized, are distributed approximately standard normal. In applied settings, psychometric indicator variables may have a skewed or bifurcated distribution which can affect the results of EFA, although such variables are usually undesirable. Another assumption of the current design is that all variables comprising each sample correlation matrix are continuous measures made on continuous latent variables, and therefore can be thought of as continuous random variables. Sometimes in applied psychometric research, however, discrete measures (such as items anchored to a Likert scale) are employed. In these cases additional methodology, such as the use of a tetrachoric or polychoric correlation matrix, may be necessary to satisfy assumptions of continuity. Finally, although the sample correlation matrices generated in this study reflect sampling error, and the addition of unique indicators introduce some random error, no error was

specifically introduced into the population matrices. Some authors (e.g., Hong, 1999; MacCallum & Tucker, 1991) have recommended the introduction of noise to the population matrix to more closely approximate a “lack of fit” between EFA models and the underlying population.

### Significance of the Study

This dissertation seeks to uncover whether, given an equal chance in direct comparison on the same data, a computational method for determining the number of factors to extract in EFA will emerge that is clearly superior to others over a wide variety of common psychometric situations. At the very least, this work seeks to discover any shortcomings of the methods under scrutiny, and to identify specific situations which may cause a procedure to fail. This dissertation also seeks to provide practical empirical evidence for those unsure of which correlation matrix to employ when attempting to solve the number of factors problem prior to analysis—a reduced or an unreduced correlation matrix. This work represents the first time this particular variable was systematically manipulated in a Monte Carlo study of this type. The eventual goal of such endeavors is to provide enough compelling evidence as to warrant the inclusion of one or more of these computational methods as pre-programmed options in future statistical packages, and to outline the situations for which they are appropriate. Thus, the factor analyst of tomorrow would have a choice other than the default eigenvalues-greater-than-one criterion from which to start a thoughtful investigation of dimensionality.

### Definition of Terms

*Communality* – The proportion of variance for an indicator variable that can be attributed to the common factors.

*Eigenvalue* – In the context of EFA, the eigenvalue represents the amount of variance in a correlation matrix that is explained by each factor or component.

*Eigenvector* – The set of orthogonal vectors each describing, in descending order, the maximal amount of variance in a correlation matrix.

*Extraction* – The method used for estimating factor loadings and communalities for a correlation matrix.

*Rotation* – A geometric transformation of the original factor solution to produce a matrix of factor loadings that is more interpretable.

*Salient* – A meaningful factor loading for an individual indicator variable.

*Uniqueness* – The proportion of variance for an indicator variable that is unique or not attributable to the common factors.



## CHAPTER II

### REVIEW OF LITERATURE

#### Factor Analysis in the Social Sciences

For decades, factor analysis has been recognized as one of the most flexible and widely used research tools in the social sciences. Once hailed as the "...reigning queen of the correlational methods" (Cattell, 1978, p. 4), its application to the field of psychological measurement and testing has been particularly significant. Even from its venerable beginnings over a century ago, factor analysis has been closely associated with the identification and disentanglement of abstract psychological constructs, such as the various facets of human intelligence (Spearman, 1904, 1927). For decades, the role of factor analysis in evaluating and refining psychological measurement instruments and in demonstrating construct validity has been well defined (American Psychological Association, 1954; Cronbach, 1971; Cronbach & Meehl, 1955; Gorsuch, 1983; Guilford, 1946; Kerlinger, 1973; Nunnally, 1978; Nunnally & Bernstein, 1994; see also Thompson & Daniel, 1996, for a discussion). Less well defined, however, has been a clear decision-making strategy vis à vis the myriad subjective choices facing the factor analyst. These decisions include, but are certainly not limited to, the choice of extraction method, the selection of an appropriate geometric rotation, and the determination of the number of non-trivial factors or components to be retained. All of these decisions have

consequences that can directly affect the outcome of factor analysis (Armstrong & Soelberg, 1968; Comrey 1978; MacCallum, 1983; Weiss, 1976).

### *Exploratory Factor Analysis versus Confirmatory*

#### *Factor Analysis*

Before jumping too quickly into a discussion of the relative minutiae of any particular factor analytic procedure, it is first appropriate to present a review of factor analytic methods from a more macroscopic perspective. To begin with, factor analysis has evolved into a somewhat loosely defined term that actually refers to a broad family of multivariate approaches to data analysis. Normally, these procedures are associated with the systematic examination of inter-item correlation and/or covariance matrices.

Subsumed within the broad category of factor analysis, then, are two main classes (Floyd & Widaman, 1995; Thompson & Daniel, 1996). The first class has been termed exploratory factor analysis (EFA); the other is confirmatory factor analysis (CFA). The former will become the focus of the present study, but the latter merits some attention here, in the interest of clarity. CFA can actually be thought of as a special type of structural equation model that focuses on proper psychometric structure. Conventionally, CFA has been most often associated with the use of LISREL-type programs, based on maximum likelihood computations (Jöreskog, 1969; Jöreskog & Sörbom, 1996).

Conceptually, the chief difference between CFA and EFA lies in whether the number of factors or components to be extracted during the procedure is known to the factor analyst a priori. Ideally, EFA would be used as a precursor to CFA, since one purpose of EFA is to provide suitable insight into the question of the number of factors to be modeled in CFA. So, having a cogent indication of the number of factors is a

prerequisite to performing CFA, the results of which then provide confirmatory evidence for the viability of the entire model (Hurley et al., 1997). During earlier stages of research, though, and throughout the initial development of a psychometric instrument, exploratory procedures are more appropriate (Hurley et al.; Thompson & Daniel, 1996). In fact, as Hurley et al. put it most succinctly:

EFA [is] appropriate for scale development while CFA would be preferred where measurement models have a well developed underlying theory for hypothesized patterns of loadings. A line of research would start out with studies utilizing EFA while later work would show what can be confirmed (p. 668).

Along these lines, a Monte Carlo study by Gerbing and Hamilton (1996) found that EFA, when used prior to cross-validation, can contribute to the improved specification of the CFA model.

### *Modes of Factor Analysis*

Historically, within the larger class of exploratory procedures, significance has been placed on the manner in which the data are collected and organized for analysis. Distinctions of this type have given rise to a taxonomy that defines six different modes of factor analysis (see Stewart, 1981, for a summary). For example, in *P*-mode factor analysis, data are collected on just one person or entity. The factors, then, are loaded across variables; but, the indices of association are computed across different occasions. Cattell and Adelson (1951) and Cattell (1953) give examples of the use of the *P* technique for analyzing changes in demographic and economic characteristics of a nation over time. In the present study, focus will be restricted exclusively to *R*-mode factor analysis. In *R*-mode, rows are cases, columns are variables, and cell entries are scores of the cases on the variables. *R*-mode is by far the most common technique of

factor analysis, so much so that it is normally assumed and not specifically labeled as such (Garson, 2007). In *R*-mode, the factors are clusters of variables on a set of people or other entities, at a given point in time. This technique permits the factor analyst to examine the relationships among items or variables for a sample of individuals.

Although somewhat different conceptually, Stewart (1981) explains, “all of the various modes of factor analysis [*P*-mode, *Q*-mode, etc.] provide information about the dimensional structure of data” (p. 52). In psychometrics, the number of factors or components extracted in EFA relate expressly to the dimensionality of a psychometric instrument. In fact, as previously mentioned, the association between factor analysis and construct validity has been long understood within the framework of classical test theory.

#### *Factor Analysis as a Tool for Establishing*

#### *Construct Validity*

Since the time of Spearman (1927) and Thurstone (1947), factor analysis has been inexorably linked to questions of validity in the psychometric literature. Guilford (1946) even makes early mention of the term “factorial validity.” He defined this term as the kind of validity that answers the question, “Does this test measure what it is supposed to measure?” (p. 428). Guilford argued that the factorial validity of a test was reflected by its loadings on meaningful, common, reference factors. Later, in an effort to codify testing standards during the 1950s, the American Psychological Association (APA) began formally defining several different types of validity in psychological measurement. The older concept of factorial validity was reintroduced under the novel name “construct validity.” The connection with factor analytical methods remained a point of emphasis, “Factor analysis is another way of organizing data about construct validity” (APA, 1954,

p. 14). Cronbach and Meehl (1955) would later explain that construct validation specifically addresses the question, “What constructs account for variance in test performance?” (p 282). The juxtaposition of psychological and statistical terms was purposeful and typical for authors such as Cronbach (APA president, 1957), who were seeking to instill more substance into the field of psychological measurement and testing. Decades later, an occasional reference to factorial validity could still be found (e.g., Nunnally, 1978, p. 87), but most mainstream nomenclature had since shifted toward terms more in line with Cronbach and the APA. Meanwhile, the close connection between construct validation and factor analytical methods never waned.

Whenever psychometric instruments are utilized in scientific research, the importance of providing empirical reliability and validity evidence simply cannot be understated. Some regard evidence of validity—especially evidence of construct validity—as fundamental to the entire discipline of psychology. Clark and Watson (1995) write, “...the process of establishing construct validity represents a key element in differentiating psychology as a science from other, nonscientific approaches to the analysis of human behavior” (p. 310). Floyd and Widaman (1995) offer some detailed examples of how exploratory factor analysis has been used to examine construct validity, and to provide insight into the multidimensional structure of particular measurement instruments.

While EFA can certainly aid in scale development, and both EFA and CFA can provide evidence for the construct validity of a certain measurement model, clean factor analytic results alone do not “prove” that a psychometric instrument is valid. As mentioned previously, it is a widely accepted tenet in the social sciences (and in other

scientific disciplines) that data can never positively confirm a model; they can only fail to disconfirm one (Popper, 1959). Similarly, finding evidence of construct validity does not confirm the present measurement model, but it is helpful in lending credence to the measurements obtained. Cronbach and Meehl (1955) explain this concept thusly, “One does not validate a test, but only a principle for making inferences” (p. 297). Clark and Watson (1995) warn against light-handed approaches to scale validation, and explicate the rigor expected within the field of psychology. They specify, “...a series of investigations is required to even begin the process of identifying the psychological construct that underlies a measure” (p. 310). It is this prior research and theory that serve to develop and define the constructs underlying psychometric instruments. Then, in the process of applied scale development, EFA is the usual mechanism for providing empirical evidence for construct validity (Kerlinger, 1973). So, use of EFA can aid in identifying, describing, and naming the factors and constructs that subsume psychological phenomena. Cliff (1983) reminds researchers, however, to guard against the nominalistic fallacy—simply naming a factor does not explain it.

### Principal Components versus Common Factor Analysis

A distinction of varying degree has historically been drawn between the method of principal components and common factor analysis. This distinction runs the gamut from fundamental to pedantic. At one end of this spectrum are theorists who regard the differences between principal components and common factor analysis as major. These theorists are divided, however, between those who would recommend principal components over the common factor model, and those who promote the opposite view.

The classic problem of factor indeterminacy (Guttman, 1955) is one inherent difference between the two methods. Schönemann (1990) reminds researchers, “Components are by definition linear combinations of the observed tests, and hence determinate. Factors, on the other hand, are only *implicitly* [italics his] defined by the factor model...it leaves the factors indeterminate” (p. 48). Incidentally, it is this vital conceptual difference—that components are linear combinations of actual measurements, while factors are not directly observable—which accounts for the pervasive use of the decidedly incommutable terms “factor” and “component” throughout the psychometric literature.

Since factor indeterminacy causes multiple problems with the computation and interpretation of individual factor scores, some see this as enough to warrant the complete abandonment of the common factor model in favor of principal components (e.g., Steiger & Schönemann, 1978). But, since psychometricians are generally more concerned with properties of the factors themselves than with individual factor scores (especially during the initial exploratory stages of test development) these concerns are not widely shared. Instead, most who emphasize the distinction between the two methods advocate the common factor model over the use of principal components. This stems from a convention within the principal components model that sets all communalities equal to one. Mathematically, this removes the error partition from the communalities, and literally implies that the model assumes all measures to be perfectly reliable. This assumption is enough to force any thoughtful psychometrician to balk, and to seriously reconsider ever employing a principal components extraction. The common factor model assumes all communalities are less than one, allowing for the modeling of error. Unfortunately in practice, communalities are unknown, and must therefore be estimated

in some way. Finding a desirable method of estimating these communalities—the so-called communality problem—has given rise to the present multitude of common factor analysis extraction techniques.

At the other extreme of the component versus factor debate are applied researchers (e.g., Velicer & Jackson, 1990a; 1990b) who point to the basic algebraic similarity between the two methods. Gorsuch (1990) puts it in simple terms, “Common factor analysis is the general case of which component analysis is a special case” (p. 33). Researchers of this ilk tend to blur the lines between the two methods. Johnson and Wichern (2002) even go so far as to refer to the method of principal components as “principal component factor analysis” (p. 517). In SAS, one may conduct a principal component analysis using PROC FACTOR, with communalities set to one. From these perspectives, factor analysis and principal component analysis are algebraic equivalents.

Bentler and Kano (1990) noted during the zenith of the raging factor versus component controversy, “Debates on the virtues of common factor analysis versus principal component analysis, and their variations, go back about 50 years to the time of Thurstone and Hotelling” (p. 67). Tracing the debate through the literature, one can clearly sense the approach of a showdown in the wake of the so many unsatisfactory attempts to solve the communality problem. As Velicer and Jackson (1990b) put it, “We view the existence of so many alternative approaches to factor analysis and the lack of a clear consensus with regard to the best method to be problematic, particularly at this stage in the development of the method” (p. 99). A statement by Wilkinson (1986) in the Version 3 manual of SYSTAT—a common desktop statistical application—that “Principal component and common factor solutions for real data rarely differ enough to



matter” sparked an at times nasty argument between applied methodologists and theoretical statisticians in the pages of *Sociological Methods and Research* (see Borgatta, 1989; Borgatta, Kercher, & Stull, 1986; Hubbard & Allen, 1987b, 1989; Wilkinson, 1989a, 1989b). Soon after, a full-blown debate would also erupt in the psychometric literature, resulting in the publication of a special issue of *Multivariate Behavioral Research* (Volume 25, Issue 1, 1990) devoted entirely to the dispute. After much clamor and a spirited exchange of opinion, theory, and empirical evidence, the editors of that classic issue gave the proponents of principal components analysis the final word. In their closing article, Velicer and Jackson (1990b) remained dissuaded from their original assertion. They insistently concluded:

1. The algebraic differences between the methods are minimal; algebraically, these differences disappear in the limit.
2. For most data sets, there will be no practical differences between the methods. (p. 110).

Velicer and his colleagues have maintained that the fundamental differences between common factor analysis and principal components analysis are theoretical and not empirical. Though the procedures can be interpreted to be quite dissimilar in theoretic terms (e.g., Widaman, 1993, p. 308), the evidence does not seem to reveal vast numeric differences. This finding is particularly pronounced when the data come from a well-designed study (Jackson & Chan, 1980; Velicer & Jackson, 1990a, 1990b; Velicer et al., 2000). Furthermore, the fact that so many alternative approaches to factor analysis exist with no clear consensus with regard to the best method is clearly problematic (Velicer & Jackson, 1990b). Even when the factor analyst does decide upon a specific common factor approach, the exact numeric values will differ depending upon choice of computer program (Wilkinson, 1989b). Principal components will at least yield values calculated in

a standard, replicable way. A point that has been conceded by those in favor of principal components, however, relates to early suspicions that principal components analysis tends to produce somewhat inflated estimates of the factor loadings (Stewart, 1981). More recent findings by Snook and Gorsuch (1989)—later replicated and extended by Widaman (1993)—have shown principal components analysis to produce a bias toward higher loadings, especially when the ratio of indicator variables to factors is low.

At this point, the consensus of the entire principal component versus common factor analysis debate appears to be that the greatest discrepancies between the methods occur when a combination of poorly identified factors and low variable saturation exist. That is, the differences are most significant when the ratio of variables per factor is very small and the loadings of the indicator variables on the factor are very low. Concerning this, Velicer and Jackson (1990b) are quick to point out, "...it should be recognized that a data set which yields poorly identified factors with low saturations is an inadequate starting point for performing either a factor analysis or a component analysis" (p. 100). In their conclusions, Velicer and Jackson go on to remind readers that although much has been written concerning the controversy surrounding the choice of factor analytic methods, perhaps this profusion of attention is unwarranted. After all, choice of methods is but one major decision facing the factor analyst; and, as it turns out, it is perhaps of the least practical importance. They warn that analysts should not fuss unduly with the relatively slight distinctions between principal components and common factor analysis, especially at the detriment of more vital decisions, such as the selection of variables and participants for inclusion in the study. It seems clear that as long as researchers exercise diligence regarding other fundamental factor analytic choices, including intelligent

variable selection, thoughtful participant sampling, and careful attention to the number of factors problem, then the choice of extraction method becomes much less pivotal.

### Importance of Rotating the Correct

#### Number of Factors

Watkins (2006) contends, “Of all the decisions made in exploratory factor analysis, determining the number of factors to extract is perhaps the most critical because incorrect specification will obscure the factor structure” (p. 344). A great deal of previous research has exposed the deleterious effects that extraction of inappropriate numbers of factors in a model can have on the results obtained (e.g., Comrey, 1978; Fava & Velicer, 1992a, 1996; Levonian & Comrey, 1966; Wood et al., 1996). For example, Fava and Velicer (1992a) demonstrated that under-extraction leads to a substantial degradation of factor scores for both principal components analysis and maximum likelihood factor analysis. While under-extraction has long been acknowledged as a non-trivial problem, over-extraction has traditionally been regarded as less serious (Thurstone, 1947; Cattell, 1978). However, more recent evidence has dispelled this notion a bit. Fava and Velicer (1992a) and Wood et al. show that over-extraction with varimax rotation can lead to factor splitting, and to the creation of false factors at the expense of true ones (see also Comrey, 1978; Lee & Comrey, 1979). For example, if data are truly unidimensional, variance from the single, real factor can be spread out over one or two false ones. Extracting an erroneous number of factors also affects the calculation of factor loading estimates. In general, under-extraction introduces much more error into the loading estimates than does over-extraction (Fava & Velicer, 1996; Wood et al.), but the factor

splitting caused by over-extraction can obfuscate the true factor structure enough to invalidate psychometric interpretation and severely hamper scale development.

## Methods for Solving the Number of

## Factors Problem

### *Eigenvalues-Greater-than-One Rule*

In a search to solve the communality problem, Guttman (1954) defined his “weakest” lower bound for the minimum rank of a population correlation matrix to be the number of characteristic roots (i.e., eigenvalues) greater than unity. As outlined in Chapter One, Kaiser (1960) extended Guttman’s result into his infamous eigenvalues-greater-than-one stopping rule for use with sample-based correlation matrices, even though Guttman specifically warns readers:

In the present paper, we do not treat the problem of ordinary sampling error, that is, of sampling a population of respondents. We assume throughout that population parameters are used, and not sample statistics. (p. 151).

Even in the face of early complaints that it overestimated the true number of factors (e.g., Browne, 1968) Kaiser would not immediately abandon the promotion of his criterion (Kaiser & Rice, 1974). Soon, though, the empirical evidence began mounting that the rule was quite likely to over-extract components (Revelle & Rocklin, 1979; Zwick & Velicer, 1982, 1986). In one Monte Carlo simulation study, Fava and Velicer (1992a) observe, “...the Kaiser eigenvalue-greater-than-one rule...typically retains approximately  $m = p/3$  components regardless of structure” (p. 395). Merenda (1997) describes Kaiser’s rule as nothing less than “one of the poorest” (p. 159) and Gorsuch (2003) even goes so far as to call it “...the prime candidate for the worst criterion ever

tried” (p. 157). Ultimately, the technique was unceremoniously rejected even by Kaiser himself (Gorsuch, 1990, p. 38).

It is now apparent that the disappointing performance of the eigenvalues-greater-than-one criterion in empirical studies is because the rule imprudently ignores the effects of sampling error. Horn (1965) demonstrated that, on average, sampling variation alone would produce eigenvalues greater than unity for half of sample eigenvalues from random data. Despite severe criticism on these grounds (e.g., Karr & Martin, 1981; Jackson, 1993), as Peres-Neto et al. (2005) recently noted, “...this method is still very popular among data analysts” (p. 980). In an effort to explain researchers’ unabated adherence to Kaiser’s imperfect criterion, Thompson and Daniel (1996) offer, “This extraction rule is the default option in most statistical packages and therefore may be the most widely used decision rule, also by default” (p. 200).

#### *Parallel Analysis*

The technique of parallel analysis (PA) was first suggested by Horn (1965) as a modification to the eigenvalues-greater-than-one rule that would take into account the effects of sampling. In PA, the researcher compares the eigenvalues calculated from the observed data to eigenvalues calculated from one or more parallel sets of random data. The random data sets are usually generated using a Monte Carlo approach, such that several matrices of random normal deviates are created with the same number of rows and columns as are in the observed data matrix. Then, eigenvalues are computed and averaged together over all the random data sets. Only those factors with observed eigenvalues greater than the corresponding eigenvalues from the random data are retained.

Some important improvements to PA have been suggested (Longman, Cota, Holden, & Fekken, 1989; see also Buja & Eyuboglu, 1992; Glorfeld, 1995). These authors recommend using a quantile (e.g., 95th percentile) of the random data eigenvalues instead of the mean; this adjustment controls for Type I error in selecting the number of factors. Buja and Eyuboglu also outline a more non-parametric implementation of PA in which random permutations of the raw data are used in lieu of the normally distributed random deviates. This modification eliminates any assumption of multivariate normality, provided that the permutations are truly random. Castellan (1992) provides a computer algorithm that guarantees the correct random shuffling of data matrices. The existence of a non-parametric implementation of PA looks to be a powerful tool in applied settings where assumptions of normality may be invalid. Notably, though, Buja and Eyuboglu report that even in the parametric instance, PA appears to be highly robust against departures from normality assumptions.

Some possible weaknesses of the PA procedure have been uncovered in the psychometric literature, though. For example, Turner (1998) and Beauducel (2001) both found that PA may underestimate the number of factors or components to retain when the first eigenvalue is large. Furthermore, Beauducel showed that PA may underestimate the number of components when a data set has oblique simple structure. Beauducel found the problem was alleviated as sample size increased, but more research on this issue is surely warranted.

As discussed in Chapter One, PA has been recommended over the eigenvalues-greater-than-one rule almost uniformly among researchers familiar with the number of factors problem. However, in 2004, Hayton et al. conducted a review of the applied

literature, looking for instances of the use of PA in the psychological subfield of industrial and organizational management. According to that research, a thorough appraisal of the *Academy of Management Journal* and the *Journal of Management* from the years 1990 to 1999 found 142 articles employing EFA techniques. Exactly zero of those articles used PA as a decision criterion for the number of factors to rotate. Results such as these raise (yet again) a question once asked by Hubbard and Allen (1987a):

Why has this test been virtually ignored by researchers? Two responses suggest themselves. First, as noted earlier, many applied researchers may be unfamiliar with alternative stopping rules and hence over rely on the [eigenvalues-greater-than-one] criterion. Second, some researchers may be reluctant to adopt a method whose implementation requires considerable time and effort.... (p 186).

Over time, various efforts have been undertaken to assuage the apparent apprehension felt by applied researchers facing a decision to employ PA. For instance, some authors (e.g., Allen & Hubbard, 1986; Longman et al., 1989; Longman, Holden, Fekken, & Xinaris, 1993; Montanelli, & Humphreys, 1976) have proposed the use of various linear extrapolations and regression equations to approximate the results of the time-consuming Monte Carlo portions of PA. Perhaps most helpful, however, have been the recent geometric improvements in computer processing speeds, coupled with well-documented public domain copies of the scripting commands for executing PA in the syntax languages of the leading statistical computing packages, such as SAS and SPSS (see O'Connor, 2000a). Also, despite early indications (e.g., Crawford & Koopman, 1973) that parallel analysis may not work in conjunction with principal axis common factoring—which involves the use of squared multiple correlations on the diagonals of the reduced correlation matrix—PA has now been successfully adapted for this purpose

(see O'Connor, 2000b). According to O'Connor, his program produces results that are equivalent to those yielded by the Montanelli and Humphreys equation. Overall, it would appear that the traditional obstacles impeding the broader use of PA are eroding, leaving few excuses for the assiduous factor analyst.

### *Scree Tests*

In Chapter One, the visual scree test (Cattell, 1966) was introduced, and weaknesses related to its historical unreliability were exposed. Recently, for example, Kanyongo (2005) found that the accuracy of the visual scree test was not reliable when used on a set of Monte Carlo generated samples from known populations. Kanyongo went on to recommend the use of MAP and/or PA in conjunction with a visual scree test. Similarly, and as mentioned previously, Velicer et al. (2000) recommend use of the visual scree only as an adjunct. Even in this role, however, some researchers may still eschew the visual scree test because, as Wood et al. (1996) point out, "...the [visual] scree test is sometimes avoided because it requires a subjective decision by the factor analyst" (p. 354). These authors prefer a more objective procedure be used to solve the delicate number of factors problem, if available.

The inherent subjectivity of the visual scree test can be further complicated by issues of scale. For example, interpretation of the visual scree can be affected by the aspect ratio of the eigenvalue plot. The third of Cattell and Vogelmann's (1977) famous four rules for interpreting a scree plot recommends that the change of the angle should be about 30° or greater, providing that ".10 on the vertical scale is drawn equal to 1.0 on the horizontal" (p. 311-312). However, as Streiner (1998) points out, "...the majority of people use the scree plots produced by computer programs, where the ratio of the axis



scales is variable and dependent on the number of factors” (p. 688). Streiner, who reported reliability estimates less than .40 in all cases, pointed to the variability of the aspect ratios in scree plots generated by standard computer packages as a possible cause. He was quick to point out that using those scree plots represented the naturalistic condition; realistically, a typical factor analyst will not take the trouble to generate a new plot with the correct aspect ratio.

Despite its inherent subjectivity, in studies of inter-rater reliability, some have found the visual scree test performs satisfactorily (e.g., Cliff, 1970; Tucker et al., 1969; Zwick & Velicer, 1982). Not surprisingly, the results improve when factor saliencies are high, sample sizes are large, and the ratio of indicator variables to factors is not low (Cliff & Hamburger, 1967; Linn, 1968). More problems arise, however, when factor structures are complex, and especially when no clear breaks in the scree are present, or when two or more apparent breaks exist (Hayton et al., 2004).

The SEscree was proposed by Zoski and Jurs (1996) as an objective counterpart to Cattell’s (1966) visual scree test. As previously mentioned, a principal criticism of the visual scree test is its inbuilt subjectivity. To address this criticism, Zoski and Jurs developed an objective method based on linear regression. Their work capitalized on guidelines set forth in Cattell (1978). Recall that when the eigenvalues from factor analysis are plotted on the ordinate of a graph, and the factor sequence is plotted on the abscissa, a “cliff” of non-trivial factors becomes distinct from the “scree” of trivial factors. Cattell specifies that the points that make up the scree should fit a line tightly. Since standard error can be thought of as a measure of the tightness of a linear fit, Zoski and Jurs recommend a systematic examination of the standard errors produced by

regressing successive points in the scree plot. These standard errors are calculated according to the usual formula for calculating the standard error of a regression line (Snedecor & Cochran, 1989, p 154):

$$s_{y \cdot x} = \sqrt{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 / (n - 2)}$$

Once the standard error  $s_{y \cdot x}$  drops below a certain threshold (specifically, the inverse of the total number of characteristic roots), the scree points have been sufficiently modeled, and the remaining points correspond to the eigenvalues of the non-trivial factors. The threshold chosen by Zoski and Jurs, which they dub the “arbiter,” is based on observations by Horn and Engstrom (1979) that the proportion of error variance in regression tends to be inversely related to sample size.

In their original article, Zoski and Jurs (1996) never explicitly mention whether they recommend using eigenvalues from a reduced or from an unreduced correlation matrix. In the comparison study by Nasser et al. (2002), from which SEscree emerged as the clear victor among regression-based variations of the visual scree test, the eigenvalues from a reduced correlation matrix were employed. Specifically, squared multiple correlations (SMCs) were placed on the diagonals of each correlation matrix, as is the practice when conducting a principal axis common factor analysis. Of course, the eigenvalues from an unreduced correlation matrix could easily be used instead, as is the case in Cattell’s (1966) original scree test design.

#### *The Minimum Average Partial*

Originally developed for use only in conjunction with unreduced correlation matrices, Velicer (1976) suggested a criterion based on the average partial correlation

between variables after removing the effects of the first  $k$  principal components. Each of these “average partials” is defined as:

$$\bar{f}_k = \sum_{\substack{i=1 \\ i \neq j}}^p \sum_{\substack{j=1 \\ i \neq j}}^p (r_{ij \cdot k})^2 / p(p-1)$$

where  $p$  is the total number of indicator variables, and  $r_{ij \cdot k}$  is the partial correlation between variables  $i$  and  $j$  when the variance associated with the first  $k$  components has been removed. Velicer suggests that the number of components that should be extracted corresponds to the number of components  $k$  that provides the smallest  $\bar{f}_k$ . Checking each average partial to find this minimum value has become known as Velicer’s minimum average partial (MAP) test. As mentioned earlier, MAP has performed exceptionally well in simulation studies; but, the previous research on this test has been conducted wholly within the framework of principal components analysis. Fabrigar et al. (1999) write that MAP is only for principal components, and has not been developed for use with common factor analysis. Similarly, Conway and Huffcutt (2003) contend, “...a limitation of [MAP] is that its use has not been extended to common factor analysis” (p. 152). However, Gorsuch (1990) writes, “...the logic of MAP can be common factor based” (p. 33). Gorsuch conceptualizes each partial correlation as a standardized index of the size of the residuals left after removing  $k$  components (Gorsuch, 2003). In order to extend MAP to work in the context of the common factor model, Gorsuch offered the following advice, “As MAP is computed from the residuals, just compute it on the residuals from [a common factor] extraction” (R. L. Gorsuch, personal communication, February 18, 2008). These residuals can be easily computed during a common factor analysis of the reduced correlation matrix.

## Use of the Number of Factors Methods in the Literature

In a landmark editorial outlining modern standards for conducting factor analytic research, the editors of *Educational and Psychological Measurement* weighed in on the question of how to determine the number of factors to extract in EFA (Thompson & Daniel, 1996). They counseled that "...more researchers should employ more sophisticated strategies for making these decisions, including parallel analysis..." (p 200). Apparently, this sound advice has thus far gone mostly unheeded in the mainstream psychometric literature. For example, in the industrial/organizational branch of psychology, where applied EFA research is quite common, high-quality EFA decisions are still apparently not. In recent years, three large-scale meta-analyses reviewed EFA articles from key journals in the industrial/organizational field (Conway & Huffcutt, 2003; Fabrigar et al., 1999; Ford et al., 1986). These three studies reviewed hundreds of articles from high-profile journals such as the *Journal of Applied Psychology*, *Personnel Psychology*, and *Organizational Behavior and Human Decision Processes* over three progressively more recent time periods. The results of these three meta-analyses are as follows. One common theme was immediately evident; in 30 to 40 percent of all the reviewed factor analytic studies, the method used to determine the number of factors went completely unreported. Given recent discussions regarding the consequences of this EFA decision, such a finding is quite alarming. On a slightly more positive note, the use of the eigenvalues-greater-than-one rule has been steadily trending down from about 22 percent to just about 15 percent since 1990. However, recall that the consensus in the psychometric literature currently calls for outright abandonment of this method. The

exclusive use of the visual scree test also seems to be on the decline, recently dropping from over 15 percent to just under 6 percent of cases. Unfortunately, only four studies were found which employed PA as the method for determining the number of factors, accounting for less than one percent of all the studies reviewed. The only really positive news from these meta-analyses is that the number of studies employing multiple criteria to solve the number of factors problem has been steadily rising over the years, from about 14 percent to almost 22 percent in 2003. Unfortunately, it is unclear how many of the studies using multiple criteria employed PA, MAP, or the SEscree test. However, these percentages do suggest that even in the mainstream psychological literature, the number of thoughtful investigations into the number of factors problem seems to be on the rise.

## CHAPTER III

### METHODOLOGY

The current study was designed to compare the accuracy, consistency, and direction of error for PA, MAP, and the SEscree procedure in determining the number of factors underlying population and sample covariance matrices generated in a Monte Carlo simulation. First, population correlation matrices with known factor structures, representing a wide variety of typical psychometric situations, were constructed and analyzed. Then, sample correlation matrices, drawn from these population correlation matrices, were simulated using the Kaiser-Dickman (1962) method. The sample correlation matrices were each analyzed by all three of the number of factors procedures. The procedures were benchmarked against one another, as well as against the well-known Kaiser-Guttman eigenvalues-greater-than-one criterion. Each procedure was evaluated for accuracy and consistency as follows. Accuracy was measured by the mean percent of times (hit rate) the procedure identified the correct number of factors. Consistency was measured by the standard error of that mean, and the direction of error was determined by examining the mean differences of any over/underestimation.

#### Monte Carlo Simulation Rationale

This study is another in a long line of Monte Carlo simulations designed to create and analyze artificial sample covariance matrices drawn from populations with known structures. Numerous factor analytic studies have been conducted through the years that

employ Monte Carlo simulation as a fundamental aspect of their methodology, and several have examined issues related to the number of factors problem (e.g., Fava & Velicer, 1992a, 1992b; Guadagnoli & Velicer, 1988; Hubbard & Allen, 1987a; Montanelli & Humphreys, 1976; Nasser et al., 2002; Tucker et al., 1969; Velicer & Fava, 1998, Velicer, Peacock & Jackson, 1982; Zwick & Velicer, 1982, 1986). Simulation studies such as these are particularly useful in the case of factor analysis, because the myriad variables of interest (e.g., sample size, number of indicators per factor, interfactor correlations, etc.) can be systematically manipulated (Hutchinson & Bandalos, 1997). Additionally, the true number of factors underlying each population covariance matrix is objectively known, since it is defined a priori and explicitly used in the computation of the population correlation matrices. Studies that evaluate the usefulness of factor analytic procedures using real data are inherently constrained to employ subjective interpretations of what might constitute the true number of factors for a given population. The ability to arbitrarily define and manipulate the different variables important to factor analysis gives researchers the control necessary to replicate a wide variety of psychometric situations that might be encountered in practice.

### Independent Variables

In the past, simulation studies exploring the number of factors question have focused on a relatively limited number of independent variable combinations. Presumably, the limited focus of previous research was because Monte Carlo studies of this nature are particularly resource intensive, and computationally time-consuming. Additionally, the programming proficiency required to facilitate a systematic manipulation of the numerous independent variables involved in a more comprehensive

study is beyond the skill of the typical factor analyst. For this dissertation, however, a fully automated SAS/IML macro was written to generate and analyze the population and sample correlation matrices of interest. This macro, executed in parallel on computers with today's faster processing speeds, allowed for a combination of independent variables to be simulated in the present study that represents an unusually comprehensive examination of assorted psychometric situations.

The levels of the independent variables included in this dissertation were chosen in accordance with the literature reviewed in Chapter Two. Even though efforts were made to include as many levels as possible, practical and theoretical considerations limited the final list of values for each independent variable to being less than exhaustive. The following is a discussion of the rationale behind each list of values for the independent variables that were manipulated in the present study.

*Number of Factors in the Population ( $m = 1, 2, 3, 4, 5, 8, 10$ )*

Many psychometric instruments that were originally envisaged by their designers to measure multiple hypothetical factors during development may turn out to only reflect a single overarching or omnibus dimension upon empirical validation (e.g., *The Nurse's Professional Values Scale*; see Leners, Roehrs, & Piccone, 2006). The ability to recognize a single factor, and distinguish it from potential noise factors, is a crucial benchmark for procedures addressing the number of factors problem. Therefore, a single factor was the lower bound for the number of factors present in the population. On the other end of the continuum, instruments with more than about 10 distinguishable, nontrivial factors are rare in applied psychometric research. Thus, 10 was used as the upper bound for the number of factors in this study. Otherwise, relatively high granularity



between 1 and 10 factors was considered important to the present design, especially from 1 to 5 factors.

*Number of Indicator Variables per Factor ( $p:m = 3, 5, 8, 10, 15$ )*

For any given population matrix in this study, an equal number of indicator variables was assumed to load on each factor. However, this ratio of items to factors was systematically manipulated from population matrix to population matrix. To ensure identifiability in a given population, the lower bound for the ratio of items to factors in factor analysis is 3:1 (Anderson & Rubin, 1956). This ratio is commonly known as the minimum identifiability constraint, and is in accordance with usual psychometric guidelines advocating the use of at least three items in the construction of any unidimensional scale ( DeVellis, 2003, chap. 3; Garson, 2007). For these reasons, no ratio of items to factors less than three was considered in the present study. It should be noted, however, that some researchers recommend using a bare minimum of four indicators per variable, based on observations of the appearance of false factors in random normal data (Humphreys, Ilgen, McGrath, & Montanelli, 1969). While it is estimated that about half of published factor analytic studies have item to factor ratios of 6:1 or higher (Fabrigar et al., 1999) instruments with more than about 15 items per factor are rare in practice. More than 15 items per factor may be undesirable for practical reasons, such as instrument length. Also, 15 well-written items should certainly be enough to saturate the domain of interest. Therefore, 15 was used as the upper bound for the number of items per factor in most cases. In a couple of extreme instances (specifically, when  $m = 8$  and when  $m = 10$ ) an upper bound of 10 was used for the number of items per factor.

*Magnitude of Factor Loadings ( $l$  = High, .6 to .8; Wide, .3 to .8; Low, .3 to .5)*

Factor loadings were allowed to vary in saturation in three specific ways. High, wide, and low ranges of factor loading coefficients were assigned following a methodology adapted from Tucker et al. (1969). For the high range of loadings, coefficients were allowed to vary in equally spaced intervals between .6 and .8. The loadings in the low range were uniformly varied from .3 to .5, while the wide range employed coefficients stretching from .3 to .8. In applied psychometric settings, factor loadings above .8 are rare. Also, items with factor loadings below .3 are normally not considered to be important indicators in practice. Horn (1969) recommends using only salient item indicators, and argues that doing so will enhance the reliability of a factor or component score. Tables 1 through 3 contain the specific variable factor loadings for each of the six  $p:m$  ratios across the high, low, and wide ranges of factor loadings.

Table 1

*Varied Saturation Factor Loadings (High, .6 to .8)*

<i>i</i>	<i>p:m</i>				
	3:1	5:1	8:1	10:1	15:1
1	0.800	0.800	0.800	0.800	0.800
2	0.700	0.750	0.771	0.778	0.786
3	0.600	0.700	0.743	0.756	0.771
4		0.650	0.714	0.733	0.757
5		0.600	0.686	0.711	0.743
6			0.657	0.689	0.729
7			0.629	0.667	0.714
8			0.600	0.644	0.700
9				0.622	0.686
10				0.600	0.671
11					0.657
12					0.643
13					0.629
14					0.614
15					0.600

*Note.* Loadings rounded to three decimal places.*i* = indicator number, *p:m* = indicator to factor ratio.

Table 2

*Varied Saturation Factor Loadings (Low, .3 to .5)*

<i>i</i>	<i>p:m</i>				
	3:1	5:1	8:1	10:1	15:1
1	0.500	0.500	0.500	0.500	0.500
2	0.400	0.450	0.471	0.478	0.486
3	0.300	0.400	0.443	0.456	0.471
4		0.350	0.414	0.433	0.457
5		0.300	0.386	0.411	0.443
6			0.357	0.389	0.429
7			0.329	0.367	0.414
8			0.300	0.344	0.400
9				0.322	0.386
10				0.300	0.371
11					0.357
12					0.343
13					0.329
14					0.314
15					0.300

*Note.* Loadings rounded to three decimal places.*i* = indicator number, *p:m* = indicator to factor ratio.

Table 3

*Varied Saturation Factor Loadings (Wide, .3 to .8)*

<i>i</i>	<i>p:m</i>				
	3:1	5:1	8:1	10:1	15:1
1	0.800	0.800	0.800	0.800	0.800
2	0.550	0.675	0.729	0.744	0.764
3	0.300	0.550	0.657	0.689	0.729
4		0.425	0.586	0.633	0.693
5		0.300	0.514	0.578	0.657
6			0.443	0.522	0.621
7			0.371	0.467	0.586
8			0.300	0.411	0.550
9				0.356	0.514
10				0.300	0.479
11					0.443
12					0.407
13					0.371
14					0.336
15					0.300

*Note.* Loadings rounded to three decimal places.*i* = indicator number, *p:m* = indicator to factor ratio.*Presence of Unique Indicator Variables (True, False)*

In a Monte Carlo simulation study comparing principal components analysis to image component analysis and maximum likelihood factor analysis, Velicer et al. (1982) augmented their population factor patterns with variables that were uncorrelated with any other variables. These “unique” indicator variables were included as a source of random error, and made for a more realistic and generalizable simulation. Although the addition of these unique variables did not have much of an effect in the Velicer et al. study, Wood et al. (1996) observed that the addition of uncorrelated variables could significantly affect the mean standard error of factor loadings when over- and under-extracting factors. Since

one goal of the present study was to simulate a wide variety of common applied psychometric situations, and since uncorrelated variables are commonly uncovered in practice, the addition of unique indicator variables was systematically manipulated. The number of unique variables augmenting the population factor pattern was set equal to the number of variables per factor in all cases. Table 4 illustrates a population target factor pattern when no unique indicator variables are present, while Table 5 is an example of the same target pattern when unique indicator variables are included.

Table 4

*Example of a Population Target Matrix with Varied Saturation and No Unique Indicator Variables (Loadings = Wide)*

$p = 25$	$m = 5$				
	1	2	3	4	5
1	0.800	0	0	0	0
2	0.675	0	0	0	0
3	0.550	0	0	0	0
4	0.425	0	0	0	0
5	0.300	0	0	0	0
6	0	0.800	0	0	0
7	0	0.675	0	0	0
8	0	0.550	0	0	0
9	0	0.425	0	0	0
10	0	0.300	0	0	0
11	0	0	0.800	0	0
12	0	0	0.675	0	0
13	0	0	0.550	0	0
14	0	0	0.425	0	0
15	0	0	0.300	0	0
16	0	0	0	0.800	0
17	0	0	0	0.675	0
18	0	0	0	0.550	0
19	0	0	0	0.425	0
20	0	0	0	0.300	0
21	0	0	0	0	0.800
22	0	0	0	0	0.675
23	0	0	0	0	0.550
24	0	0	0	0	0.425
25	0	0	0	0	0.300

*Note.* Indicator to factor ratio  $(p:m) = 5$ .  $p$  = number of indicators,  $m$  = number of factors.

Table 5

*Example of a Population Target Matrix with Varied Saturation and Unique Indicator Variables Present (Loadings = Wide)*

$p' = 30$	$m = 5$				
	1	2	3	4	5
1	0.800	0	0	0	0
2	0.675	0	0	0	0
3	0.550	0	0	0	0
4	0.425	0	0	0	0
5	0.300	0	0	0	0
6	0	0.800	0	0	0
7	0	0.675	0	0	0
8	0	0.550	0	0	0
9	0	0.425	0	0	0
10	0	0.300	0	0	0
11	0	0	0.800	0	0
12	0	0	0.675	0	0
13	0	0	0.550	0	0
14	0	0	0.425	0	0
15	0	0	0.300	0	0
16	0	0	0	0.800	0
17	0	0	0	0.675	0
18	0	0	0	0.550	0
19	0	0	0	0.425	0
20	0	0	0	0.300	0
21	0	0	0	0	0.800
22	0	0	0	0	0.675
23	0	0	0	0	0.550
24	0	0	0	0	0.425
25	0	0	0	0	0.300
26	0	0	0	0	0
27	0	0	0	0	0
28	0	0	0	0	0
29	0	0	0	0	0
30	0	0	0	0	0

*Note.* Indicator to factor ratio  $(p:m) = 5$ .  $p'$  = total number of indicators,

$m$  = number of factors.



*Magnitude of Interfactor Correlations ( $r = 0, .2, .4$ )*

An interfactor correlation of zero represents the case where factors in the underlying population are not correlated with one another, and is the true lower limit for an interfactor correlation. This population structure would manifest itself most clearly after an orthogonal rotation of the original factor solution. Typically one would use a varimax rotation (Kaiser, 1956, 1958) when the underlying factor structure is thought to be orthogonal. Oblique rotations of the factor solution are employed to reveal simple structure when correlations between the factors are suspected. Correlations between factors may have an effect on the performance of computational methods for determining the number of factors. For example, Beauducel (2001) found that the method of PA tended to underestimate the true number of factors when dealing with data sets that had oblique simple structure and more than eight factors.

In applied psychometric settings, some correlation between factors is common and expected. High correlations, however, are problematic. When factors begin to correlate with one another too highly, it may be an indication that the factors are not practically distinguishable phenomena. Therefore, .4 was set as the upper limit for interfactor correlations in this study. The values between the upper and lower limits were chosen to provide reasonable granularity along this continuum.

*Sample Size ( $n = 250, 500, 1000$ )*

The size of each random sample drawn from a population was fixed at 250, 500, or 1000; as such, this independent variable was not completely crossed with all levels of the other independent variables. Instead, the sample size was set to 250 in situations with fewer factors and indicators, and the sample size was set to 1000 in situations with more

factors and indicators—based on traditional guidelines in the psychometric literature recommending sample sizes exceeding 5 to 10 times the number of indicators (Comrey, 1978; Comrey & Lee, 1992; Gorsuch, 1983; Guilford, 1946). Additionally, however, sample sizes were adjusted in accordance with recommendations by Mundfrom, Shaw, and Ke (2005), who showed that sample sizes should be larger when the  $p:m$  ratio is low, but can be smaller when the ratio is high. The three sample sizes correspond to small, moderate, and large sized studies respectively, and were considered adequate to produce a stable and replicable factor solution in all cases. Table 6 summarizes the specific design conditions used in the current study.

*Principal Components versus Common Factors (Reduced, Unreduced)*

As explained in previous sections, the choice of whether to use principal components or to apply the common factor model when examining the number of factors problem condenses to a choice between using the reduced or the unreduced correlation matrix when computing the eigenvalues and residuals necessary as inputs to the computational procedures. From a pragmatic standpoint, the choice which yields the most accurate recovery of the dimensionality of the population target pattern will be considered the most appropriate, regardless of the model ultimately used in the analysis.

Table 6  
*Design Conditions for Sample Data*

$n$	$p$	$p'$	$p:m$	$n$	$p$	$p'$	$p:m$
$m = 1$				$m = 2$			
250	3	6	3:1	250	6	9	3:1
250	5	10	5:1	250	10	15	5:1
250	8	16	8:1	250	16	24	8:1
250	10	20	10:1	250	20	30	10:1
250	15	30	15:1	250	30	45	15:1
$m = 3$				$m = 4$			
500	9	12	3:1	500	12	15	3:1
500	15	20	5:1	500	20	25	5:1
500	24	32	8:1	500	32	40	8:1
500	30	40	10:1	500	40	50	10:1
500	45	60	15:1	500	60	75	15:1
$m = 5$				$m = 8$			
1000	15	18	3:1	1000	18	21	3:1
500	25	30	5:1	500	40	45	5:1
500	40	48	8:1	500	64	72	8:1
500	50	60	10:1	500	80	90	10:1
500	75	90	15:1				
$m = 10$							
1000	30	33	3:1				
500	50	55	5:1				
500	80	88	8:1				
500	100	110	10:1				

*Note.*  $n$  = sample size,  $p$  = number of indicators,  $p'$  = total number of indicators when unique indicators are present, and  $m$  = number of factors.

The unreduced correlation matrix has unities in the diagonal. In the case of a reduced correlation matrix, the unities in the diagonal are replaced with communality estimates less than unity. The estimation of these communalities corresponds directly to

which flavor of common factor analysis is being employed for the analysis. In this dissertation, SMCs will be inserted into the diagonals of the correlation matrix prior to the use of a principal axis program to compute eigenvalues and residuals. This very specific method of common factor analysis is called principal axis factoring, and is the same variety of common factor model used in the Nasser et al. (2002) study and the Humphreys and Illgen (1969) paper. Although Crawford and Koopman (1973) warn that PA may not work in conjunction with principal axis factoring, O'Connor (2000b) claims that his public domain software—which allows for PA of both the unreduced correlation matrix and the reduced correlation matrix with SMCs on the diagonals—has produced results essentially identical to those yielded through the use of the Montanelli and Humphreys (1976) equation. Clearly, more empirical investigation of this issue is needed.

#### Other Design Specifications

The number of sample correlation matrices generated per condition was held constant at 100 replications. This number was considered sufficient for a resampling design, and is the same number of sample correlation matrices used by Nasser (1997) and Beauducel (2001). The number of population iterations for PA was set to 50. Longman and Holden (1992) suggest using at least 40 permutations of the raw data to provide good estimates of the mean eigenvalues. The percentile cutoff for PA was fixed at 95 percent. Setting the percentile cutoff for PA to 95 percent is equivalent to setting a 5 percent significance level for the inclusion of a factor (Buja & Eyuboglu, 1992). These authors recommend setting the cutoff for PA to 95 percent. When data are suspected to deviate greatly from normality, the nonparametric form of PA should be implemented, in which random permutations of the raw data matrix are employed. PA is easier to implement,

however, and less computationally intensive in its parametric form, which utilizes only random normal deviates. Interestingly, Buja and Eyuboglu, who first described the non-parametric implementation of PA, also found that the parametric form of PA is quite robust against departures from assumptions of normality. The probability distribution of the indicator variables in each sample was assumed to be standard normal. Although encountered in practice, variables with distributions deviating severely from assumptions of normality were considered to be beyond the scope of this dissertation.

## Summary of the Study Design

### *Independent Variables*

1. Number of factors in the population (1, 2, 3, 4, 5, 8, 10).
2. Number of indicators (items) per factor – not completely crossed (3, 5, 8, 10, 15).
3. Magnitude of factor loadings (High, .6 to .8; Wide, .3 to .8; Low, .3 to .5).
4. Presence of unique indicator variables (True, False).
5. Magnitude of interfactor correlations (0, .2, .4).
6. Sample Size – not completely crossed (250, 500, 1000).
7. Principal components versus common factors (Reduced, Unreduced).

### *Other Design Specifications*

1. Number of sample correlation matrices generated per condition (100).
2. Number of population iterations for PA (50).
3. Percentile cutoff for PA (95th).
4. Form of PA (Parametric, Random Normal Deviates).
5. Probability distribution of the variables in each sample (Standard Normal).

Completely-crossed  $m$  conditions:  $(5 \text{ levels of } m) \times (5 \text{ levels of } p:m)$ .

Extreme  $m$  conditions:  $(2 \text{ levels of } m) \times (4 \text{ levels of } p:m)$ .

Factor loading conditions:  $(3 \text{ levels of } l)$ .

Unique indicator conditions:  $(2 \text{ levels})$ .

Interfactor correlation conditions:  $(3 \text{ levels of } r)$ .

Eigenvalue conditions:  $(2 \text{ levels})$ .

Total number of design conditions:  $(5 \times 5 + 2 \times 4) \times 3 \times 2 \times 3 \times 2 = 1188$

### Computer Program

To empirically investigate the six research questions outlined in Chapter One, a SAS/IML software macro was coded in SAS version 9.1.3 for Windows. An annotated copy of the entire program can be found in Appendix A. Portions of the code have been adapted from open source code for conducting PA, MAP, and the SEscree procedure. The code for conducting PA and MAP was based on programs published by O'Connor (2000a). However, modifications to this original code were made based on software found in the public domain (see "SAS program for determining the number of components using parallel analysis," 2006; "SAS program for determining the number of components using Velicer's MAP test," 2006). The MAP procedure was adapted to be based on the residuals from either a reduced or an unreduced correlation matrix according to guidelines set forth by R. L. Gorsuch (personal communication, February 18, 2008) as described in Chapter Two. The code for conducting the SEscree procedure as well as parts of the data generation subroutine was adapted from scripted SAS commands found in Nasser (1997) and Nasser et al. (2002).

### Data Generation

In 1962, Kaiser and Dickman developed a method for generating sample correlation matrices from a given population matrix using the fundamental postulate of component analysis. Let  $\mathbf{Z}_{p \times n}$  be the population score matrix. Then,

$$\mathbf{Z}_{p \times n} = \mathbf{F}_{p \times p} \mathbf{X}_{p \times n}$$

where  $\mathbf{F}_{p \times p}$  is a principal components factoring of the positive definite population correlation matrix  $\mathbf{R}_{p \times p}$  and  $\mathbf{X}_{p \times n}$  is a population score matrix on the components represented in  $\mathbf{F}_{p \times p}$ . To empirically investigate the research questions outlined in

Chapter One, a SAS/IML macro was coded utilizing this relationship to generate population and sample correlation matrices by the Kaiser-Dickman method.

For each combination of the independent variables summarized previously, the following procedure was applied. First, a population target pattern matrix  $\mathbf{P}_{p \times m}$  was constructed according to the values specified by each independent variable. Then, the population correlation matrix  $\mathbf{R}_{p \times p}$  was computed, based on the common factor model (Gorsuch, 1983). This computation utilized a matrix of uniform interfactor correlations  $\mathbf{C}_{m \times m}$  as well as a diagonal uniqueness matrix  $\mathbf{U}_{p \times p}$  such that

$$\mathbf{R}_{p \times p} = \mathbf{P}_{p \times m} \mathbf{C}_{m \times m} \mathbf{P}_{m \times p}' + \mathbf{U}_{p \times p}$$

The values in  $\mathbf{U}_{p \times p}$  were calculated using the relationship

$$u_j^2 = 1 - h_j^2,$$

where  $h_j^2$  ( $j = 1, 2, \dots, p$ ) was the communality of each indicator. In the population, these communalities have been shown to be equal to the squared multiple correlation (Guttman, 1940). Application of this relationship guaranteed that unities were on the diagonals of every  $\mathbf{R}_{p \times p}$ .

Next, from each population correlation matrix  $\mathbf{R}_{p \times p}$ , a population principal components pattern matrix  $\mathbf{F}_{p \times p}$  was derived. To generate each sample data matrix  $\hat{\mathbf{Z}}_{p \times n}$  the population principal components pattern matrix  $\mathbf{F}_{p \times p}$  was multiplied by a matrix of random standard normal deviates  $\hat{\mathbf{X}}_{p \times n}$  as follows:

$$\hat{\mathbf{Z}}_{p \times n} = \mathbf{F}_{p \times p} \hat{\mathbf{X}}_{p \times n}$$



To be specific, the random standard normal deviates generated by SAS are pseudo-random standard normal numbers based on a Box-Muller transformation of the UNIFORM function (Box & Muller, 1958).

Finally, each sample data matrix  $\hat{\mathbf{Z}}_{p \times n}$  was used to compute each sample correlation matrix  $\hat{\mathbf{R}}_{p \times p}$ . These sample correlation matrices were used as the main input for each of the three computational methods for determining the number of factors to extract. In the special case of non-parametric PA, the sample data matrix  $\hat{\mathbf{Z}}_{p \times n}$  can be used directly.

### Data Analysis

To evaluate the performance of the three computational rules, 100 random sample correlation matrices were generated for each cell of the experimental design, drawn from a population correlation matrix constructed as described above. So, for each combination of the independent variables under consideration in the present design, the three computational methods were employed to determine a solution to the number of factors problem for each of the 100 sample correlation matrices, as well as for the population correlation matrix. Additionally, the eigenvalues-greater-than-one rule was applied to each of the sample correlation matrices, and to each population correlation matrix.

Since each target population pattern was constructed a priori, the true or target number of factors underlying the population correlation matrix was known a priori as well. The accuracy of each of the computational procedures and of the eigenvalues-greater-than-one rule was measured by the mean percent of times the procedure or rule identified the target number of factors (i.e., the mean hit rate). As previously mentioned, consistency was measured by the standard error of the mean hit rate, and the direction of

error, or tendency to over- or under-extract, was measured by mean differences. The measures were compared across the conditions specified above, and applied to each of the six empirical research questions under investigation in this dissertation.

## CHAPTER IV

### RESULTS

In all, 1188 population matrices were generated for this study, each satisfying a specific condition as outlined in Chapter Three. From each of the population matrices, 100 sample matrices were drawn, and the number of accurate hits were counted for each of the number of factors procedures, as well as the size and direction of any difference. For a given population condition, the hit rates and differences were taken directly. For a given sample condition, the hit rates and differences were first averaged across the 100 samples. Table 7 shows the overall accuracy of each of the computational procedures under investigation here, over all of the various psychometric conditions outlined above.

#### Population Results

##### *Standard Error Scree Test*

In line with Nasser et al. (2002), the common factors-based version of the SEscree appeared to have little trouble identifying the correct number of factors present in population correlation matrices, with a mean hit rate of over 92 percent, and a tendency to just slightly over-factor. When using eigenvalues from an unreduced population correlation matrix, however, the accuracy of the SEscree declined sharply to 46 percent. Additionally, the mean discrepancy between the true number of factors present in the population and the number recommended by SEscree approached ten factors too many.

Table 7

*Overall Accuracy of the Number of Factors Methods*

Matrix	SEscree		MAP		PA		KG	
	Hit Rate	Diff	Hit Rate	Diff	Hit Rate	Diff	Hit Rate	Diff
Population								
Unreduced	46.0%	9.61	31.3%	6.71	100.0%	0.00	100.0%	0.00
Reduced	92.1%	0.42	100.0%	0.00	100.0%	0.00	68.7%	-1.35
Sample								
Unreduced	38.6%	5.33	66.4%	-1.43	91.7%	-0.10	23.2%	5.87
Reduced	44.7%	4.42	4.6%	15.02	83.3%	0.19	68.2%	-1.09

*Note.* Accuracy was measured over all conditions,  $n = 594$  for each mean hit rate and for each mean difference. A positive mean difference indicates a tendency to over-factor, while negative mean differences are indicative of under-factoring.

*Minimum Average Partial Test*

The common factors version of MAP was 100 percent accurate when examining population matrices, while its principal components-based analogue was much less accurate. Principal components-based MAP correctly identified the true number of factors present in population matrices only about 30 percent of the time, with a mean overestimation of almost seven factors.

*Parallel Analysis*

PA was flawlessly accurate at detecting the correct number of factors present in population correlation matrices, regardless of whether eigenvalues from the reduced or the unreduced population correlation matrix were used for the analysis.

### *Eigenvalues-Greater-than-One Rule*

As a baseline, the Kaiser (1960) criterion was applied to all eigenvalues generated in the present work. Not surprisingly, in population matrices, the number of principal components eigenvalues greater than one always corresponded to the true number of factors present in the population. Using eigenvalues from a principal axis factoring of the population matrix, however, hampered accuracy. Hit rates in this situation dropped to about 69 percent, although the mean difference was only -1.35 factors.

### Sample Results

#### *Standard Error Scree Test*

As shown in Figure 7, the promising common factors-based version of the SEscree described by Nasser et al. (2002) failed to identify the correct number of factors more than half of the time. When the procedure was incorrect, it tended to recommend over-factoring by more than four factors on average. These findings are far less impressive than the Nasser et al. results might have implied. As a separate replication check, a single sample matrix from each of the 120 population conditions described by Nasser et al. was evaluated using the software developed for this study. In every case, the common factors-based SEscree procedure retained a number of factors within a single standard deviation of the published means (Nasser et al. p. 406-409). It appears, then, that the performance of the common factors-based SEscree procedure degrades with the addition of more realistic and extended study conditions, such as those in the present design. Unfortunately, the use of eigenvalues from an unreduced correlation matrix did nothing to improve the accuracy of this procedure. The hit rate for this principal components-based incarnation of the SEscree was an unimpressive 38.6 percent, and the

technique exhibited a tendency to overestimate the true number of factors by more than five.

#### *Minimum Average Partial Test*

When using principal components-based eigenvalues as intended by Velicer (1976), MAP was found to be highly accurate across various psychometric conditions. The original version of MAP correctly identified the true number of factors in sample correlation matrices almost two-thirds of the time, and tended to under-extract by an average of only 1.43 factors. However, across the various study conditions, the common factors version of MAP performed the most poorly of all the methods under examination here. This method, which was flawlessly accurate when used to examine population correlation matrices, simply collapsed when applied to sample correlation matrices.

#### *Parallel Analysis*

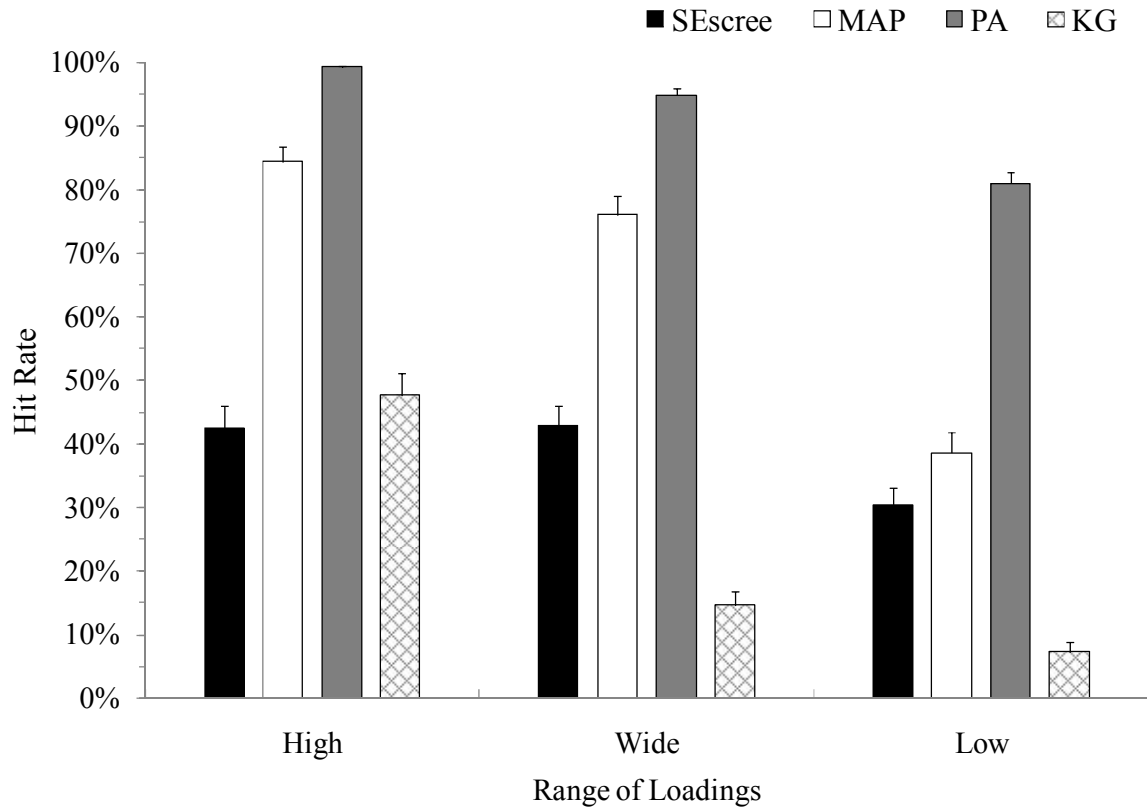
Of all the methods under examination in the present study, the most accurate method was clearly the principal components-based version of PA, which hit the correct number of factors an impressive 91.7 percent of the time across all study conditions. When it did miss, this procedure was only off by an average of -.10 factors. Although PA has been demonstrated to outperform other procedures time and again—as detailed in Chapter One—this level of accuracy across so many combinations of psychometric conditions was surprising. Almost as impressively, the common factors-based analogue of PA was observed to be dead-on accurate over 83 percent of the time, and tended to only slightly over-factor (mean difference = 0.19).

### *Eigenvalues-Greater-than-One Rule*

The Little Jiffy approach (Kaiser, 1970) which advises the use of principal components-based characteristic roots evaluated according to the eigenvalues-greater-than-one criterion performed exactly as indicated by the psychometric literature. As mentioned above, this strategy performed flawlessly for population matrices, but incorrectly identified the true number of factors in sample correlation matrices more than three-quarters of the time. On average, Little Jiffy recommended the extraction of almost six extra factors when incorrect. Using the Kaiser criterion to evaluate eigenvalues calculated from the reduced sample correlation matrices proved more successful. The mean hit rate using this approach was near 68 percent, and actually resulted in some slight under-extraction, on average.

### *Effects of Specific Design Conditions*

In the section that follows, the overall results outlined above will be broken down in more detail, controlling for specific design conditions. The impact of specific target conditions on performance will be examined, one by one. The effects of the different ranges of factor loadings, the number of factors and indicator variables present in the target population, the various interfactor correlations, and the presence or absence of unique indicator variables will be individually considered. For each comparison, the results will be presented separately for the principal components-based methods and for the common factors-based methods.



*Figure 1.* Mean hit rates (+SE) over all conditions by range of factor loadings (principal components-based methods).

Figure 1 and Figure 2 show the effects of the three different loading conditions on the accuracy of the number of factors procedures. As expected, lower salient factor pattern coefficients had a detrimental effect on accuracy for all of the methods. As shown in Figure 1, the principal components-based version of PA performed almost flawlessly over all conditions when the loading range was high, and remained highly accurate when the loading range was wide. The common factors-based version of PA appeared to be less resilient than its principal components-based cousin when the range of factor loadings was wide and when it was low.



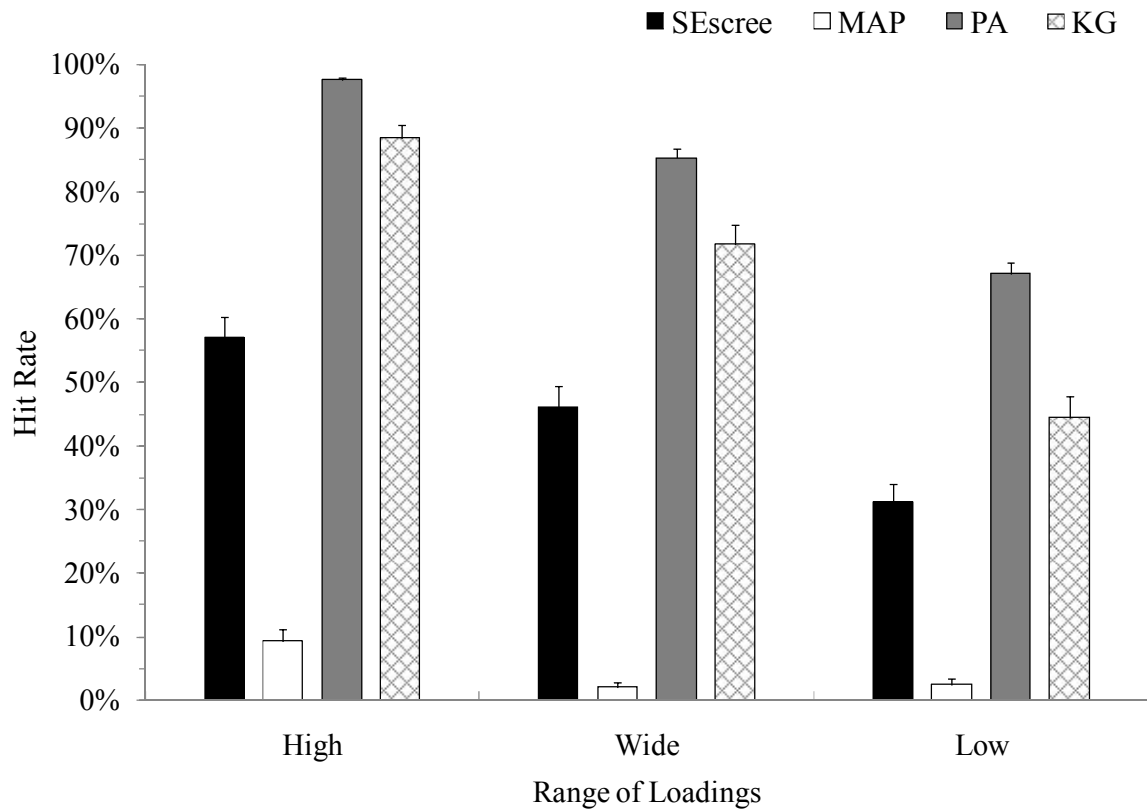
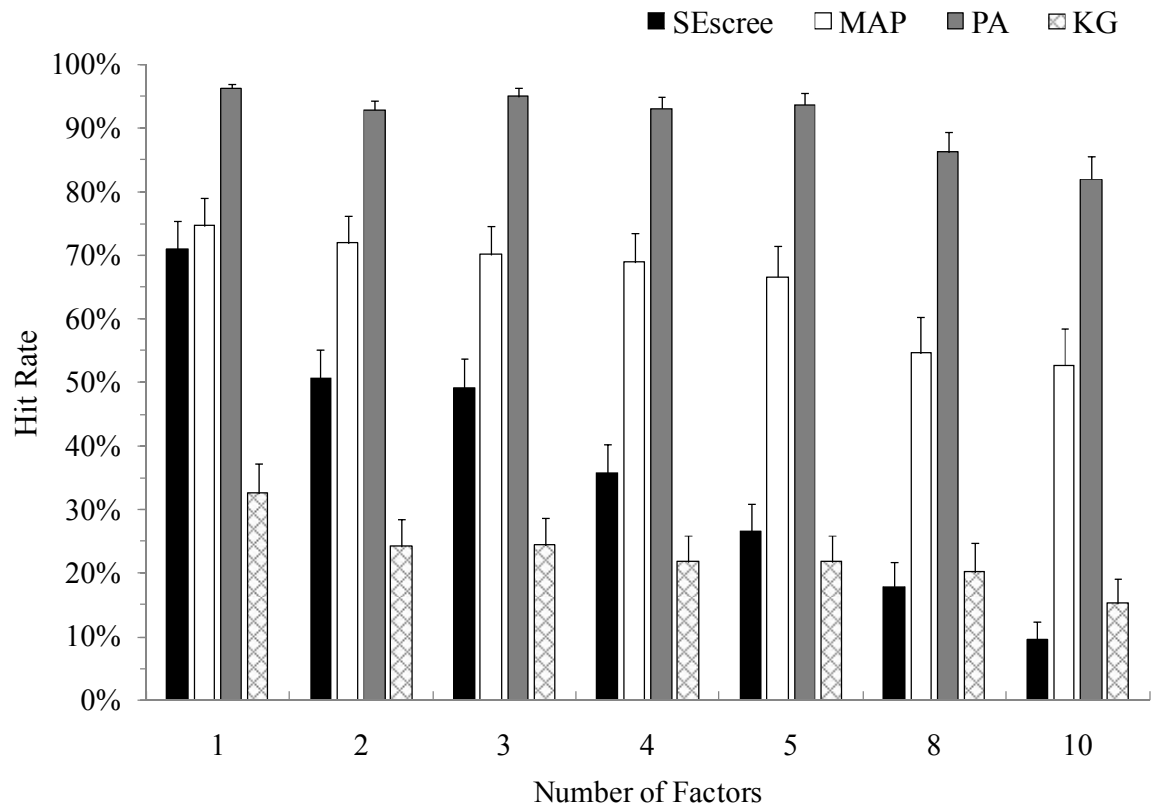
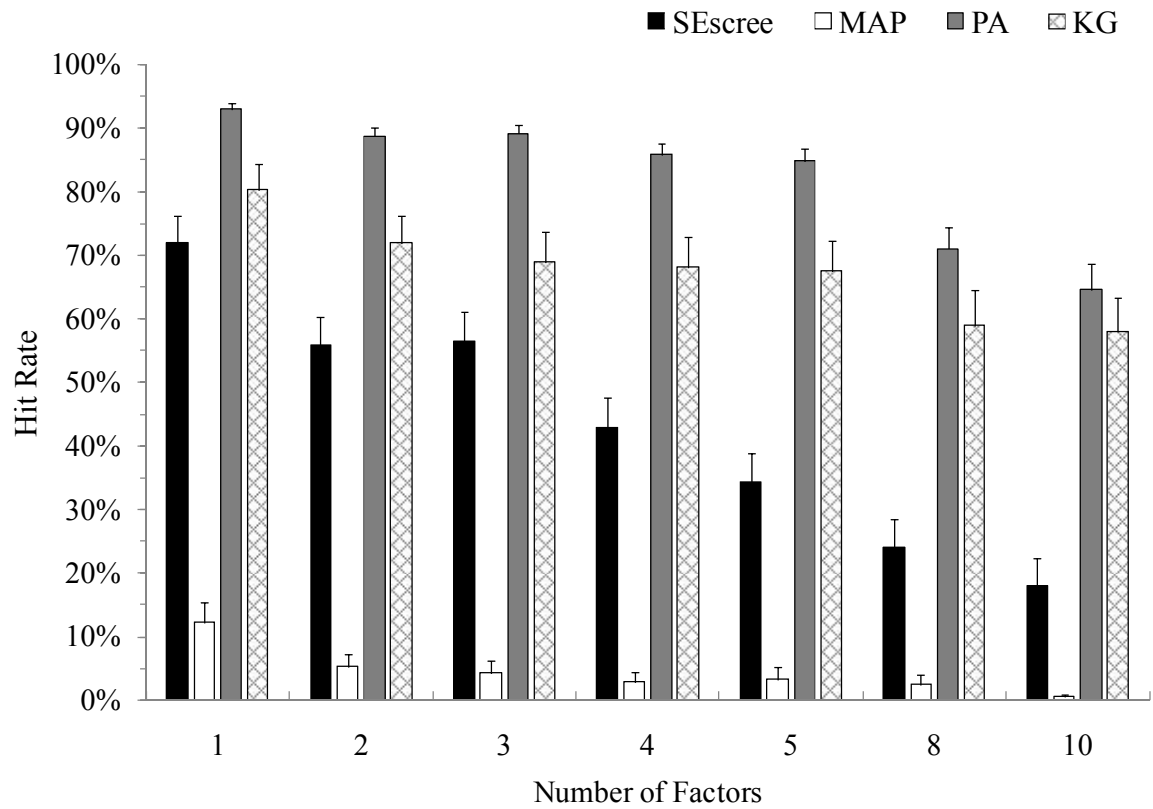


Figure 2. Mean hit rates (+SE) over all conditions by range of factor loadings (common factors-based methods).

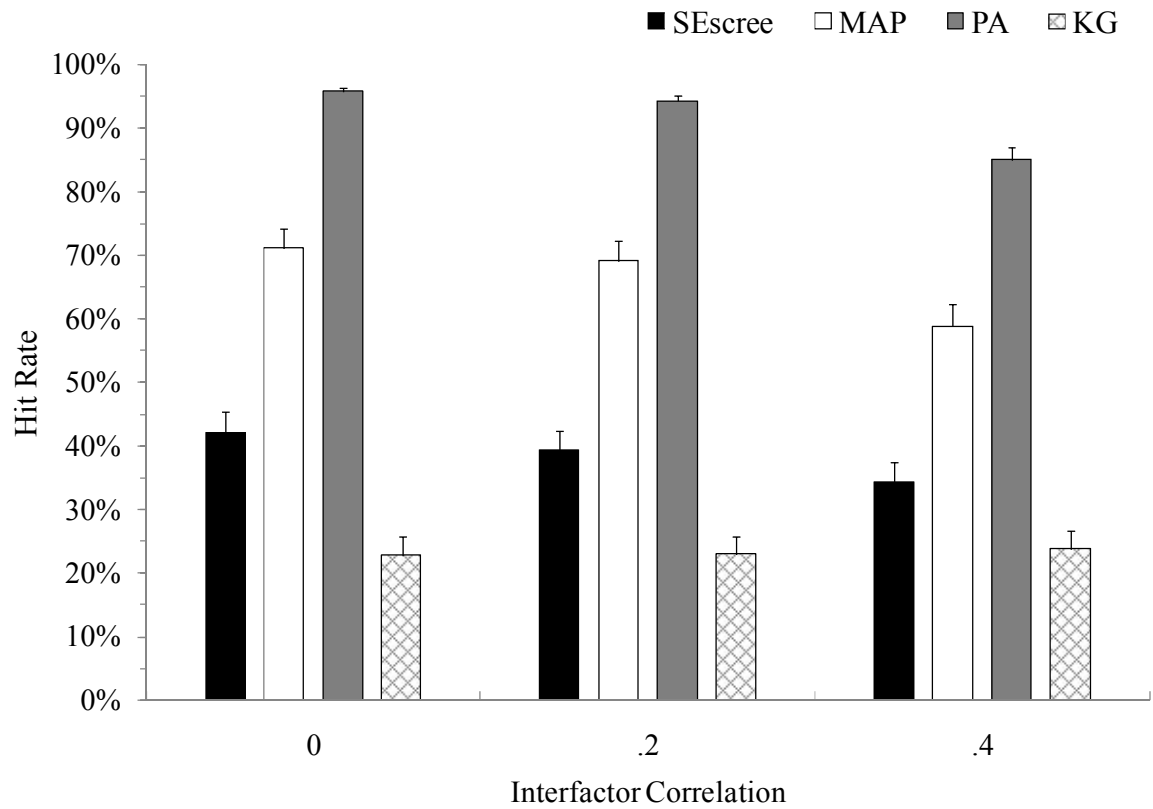
As depicted in Figure 3 and Figure 4, the performance of each method for solving the number of factors problem declines as the number of factors present in the target population increases. The SEscree procedure appears particularly susceptible to this phenomenon, and its performance degrades rapidly as the number of factors present increases. Overall, the principal components-based version of PA performs at a high level, even when the number of factors present in the population reaches ten. At this extreme, PA still achieves a hit rate of more than 80 percent.



*Figure 3.* Mean hit rates (+SE) across all conditions by number of factors (principal components-based methods).



*Figure 4.* Mean hit rates (+SE) across all conditions by number of factors (common factors-based methods).



*Figure 5.* Mean hit rates (+SE) across all conditions by interfactor correlations (principal components-based methods).

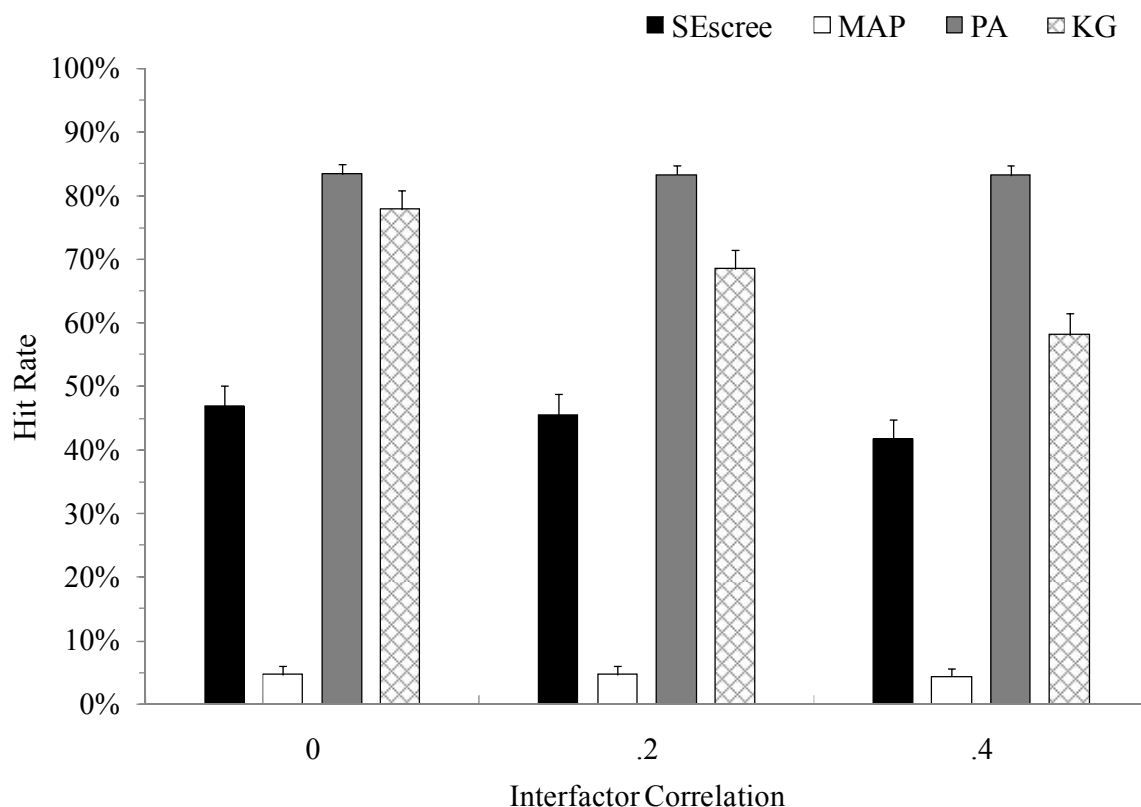


Figure 6. Mean hit rates (+SE) across all conditions by interfactor correlations (common factors-based methods).

The effects on performance of each of the number of factors methods as interfactor correlations become increasingly oblique are illustrated in Figure 5 and in Figure 6. Performance appears reasonably stable across methods as long as interfactor correlations are low, but degrades somewhat when factors become more highly correlated. Recall that Beauducel (2001) found principal components-based PA tended to underestimate the true number of factors when dealing with data sets that had oblique simple structure and more than eight factors. Table 8 shows how this finding was replicated in the present study. Hit rates for the PCA version of PA were observed to drop to approximately 65 percent when the number of factors and the magnitude of interfactor

correlations were both large. Additionally, when performance was observed to decline, the method typically erred toward under-extraction, in agreement with the Beauducel result. The common factors-based variations of SEscree, MAP, and PA, however, appear to be more impervious to degradation, with PA far outperforming the others, but not quite reaching the levels of accuracy seen in its principal components-based cousin.

Table 8

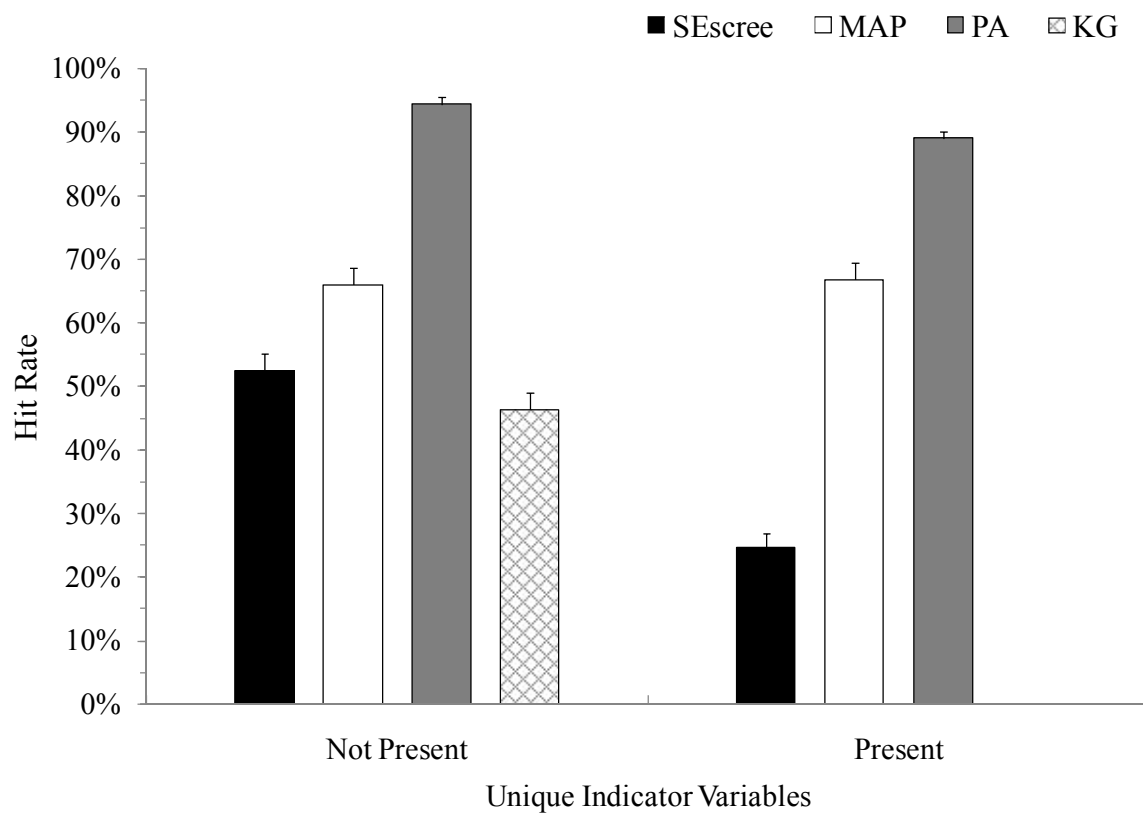
*Accuracy of PA when the Number of Factors is Large*

$r = 0$		$r = .2$		$r = .4$	
Hit Rate	Diff	Hit Rate	Diff	Hit Rate	Diff
$m = 8$					
94.4%	0.06	91.5%	-0.02	72.7%	-0.79
$m = 10$					
92.8%	0.08	87.5%	-0.07	65.6%	-1.32

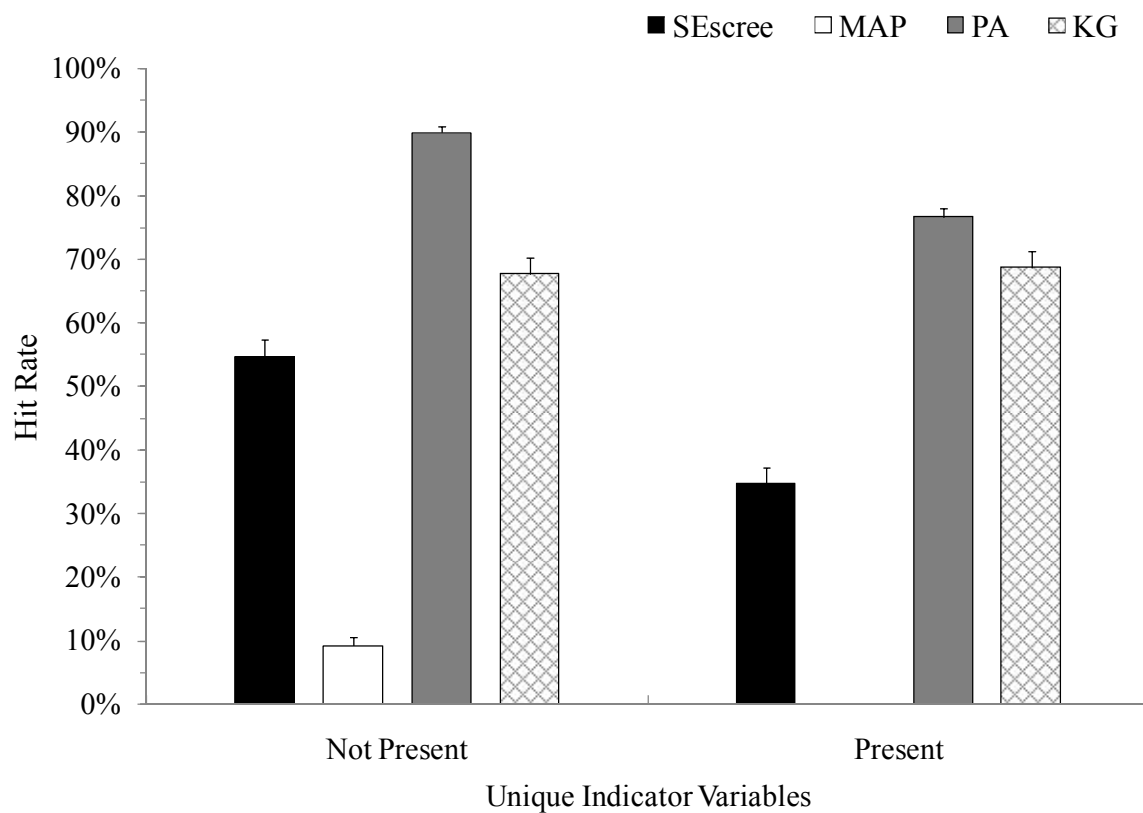
*Note.* Values are for the principal components-based version of PA.

$m$  = number of factors;  $r$  = interfactor correlations.

The effect of introducing uncorrelated or unique indicator variables into the population correlation matrix is depicted in Figure 7 and in Figure 8. The addition of these variables appears to be generally detrimental, adversely affecting performance in most cases. Notable exceptions to this generality are the principal components variants of MAP and PA. The common factors-based application of the eigenvalues-greater-than-one rule also appears resistant to the addition of these noise factors.

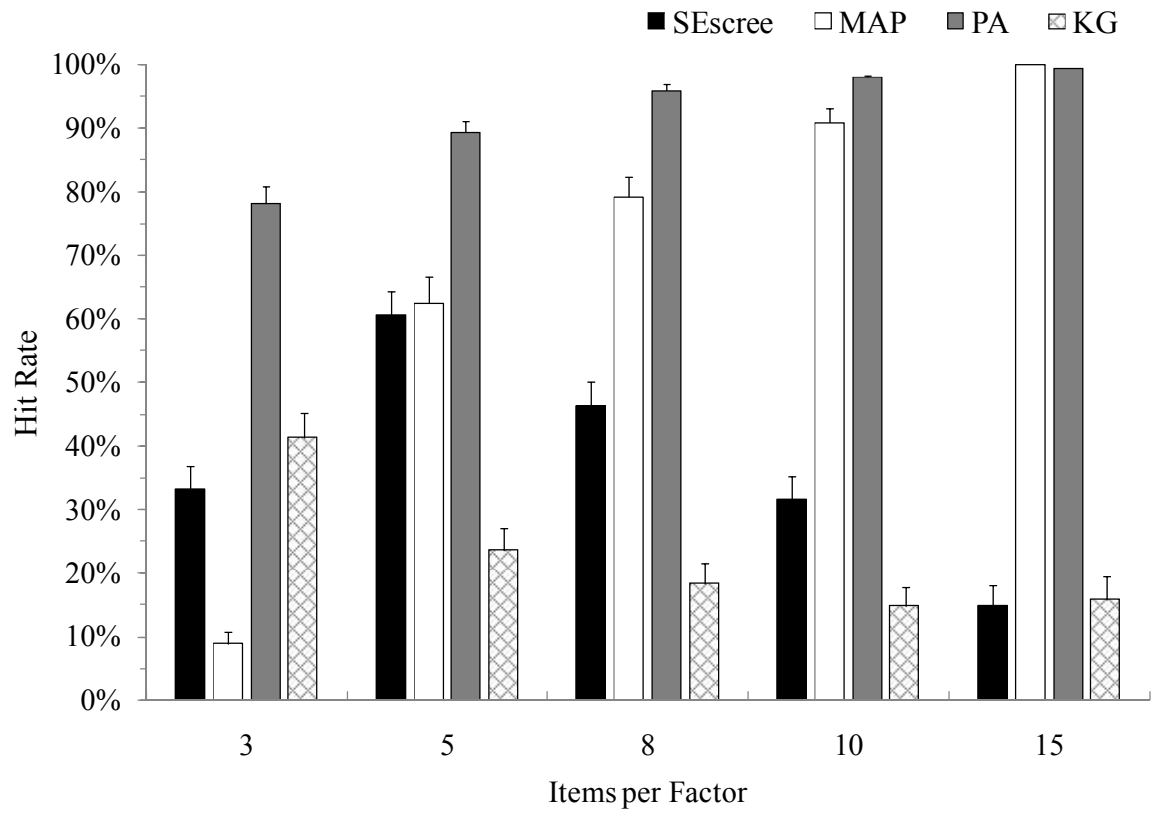


*Figure 7.* Mean hit rates (+SE) across all conditions by presence of unique indicator variables (principal components-based methods).

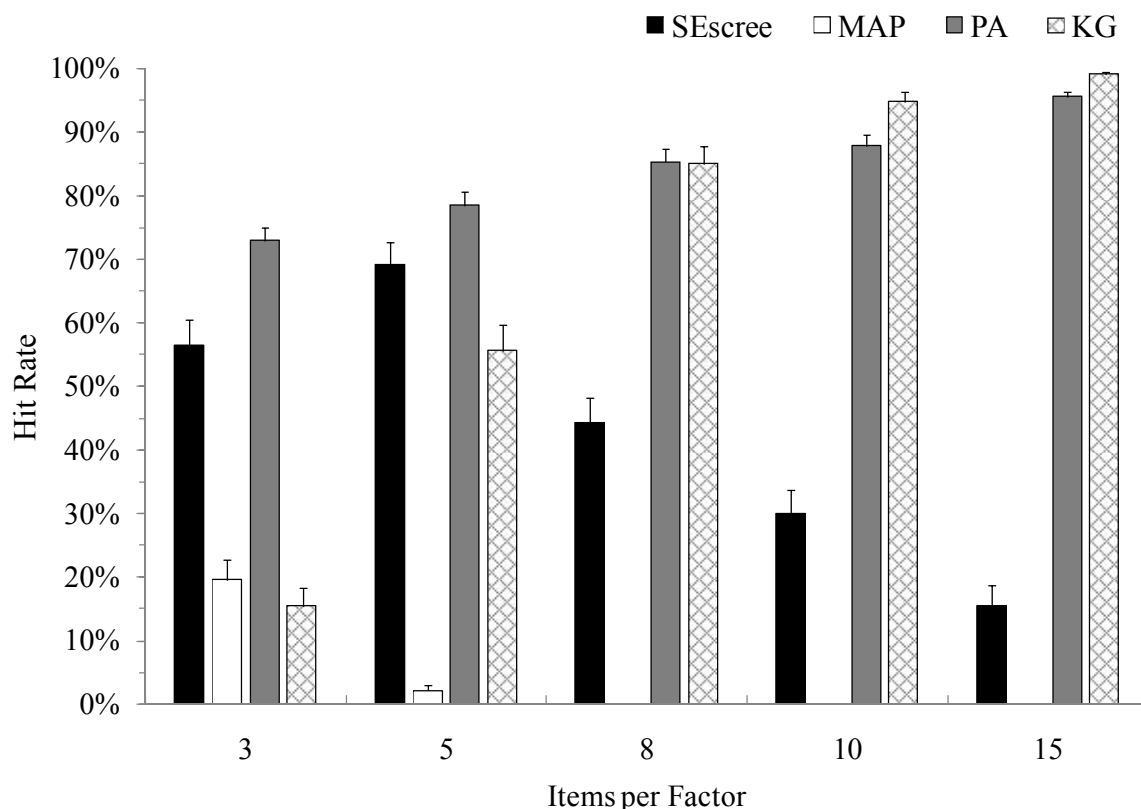


*Figure 8.* Mean hit rates (+SE) across all conditions by presence of unique indicator variables (common factors-based methods).





*Figure 9.* Mean hit rates (+SE) across all conditions by number of items per factor (principal components-based methods).



*Figure 10.* Mean hit rates (+SE) across all conditions by number of items per factor (common factors-based methods).

Figure 9 and Figure 10 give some insight into the cause of the disappointing overall performance results observed for the promising SEscree procedure. As the number of items per factor increases, the accuracy of the methods was expected to increase. This was indeed the case for PA and for the principal components version of MAP; however, this was decidedly not the case for SEscree. Clearly, more items per factor hurt the performance of SEscree. This finding is actually consistent with tabled values found in Nasser et al. (2002, p. 406) which hint that SEscree may begin to severely over-factor when the total number of indicator variables is large. This problem is likely due to problems with the “arbiter” threshold suggested by Zoski and Jurs (1996),

which is calculated as a function of the total number of characteristic roots present in the data.

Special attention was given to the specific situation when the number of factors was high, the interfactor correlations were high, and the range of salient factor loadings was low. This combination of events presented the steepest challenge to all of the methods, and was the one area where the performance of principal components-based PA appeared to falter. Given this specific combination of events ( $m = 10$ ;  $r = .4$ ;  $l = \text{Low}$ ), principal components-based PA was observed to under-factor by an average of 3.27 factors. Interestingly, even in this troublesome situation, PA actually found the correct number of factors over 28 percent of the time, while each of the other methods was observed to completely blow up, with hits rates below three percent, and mean differences ranging from -9 to +41.

## CHAPTER V

### DISCUSSION

The purpose of this study was to directly compare the three leading computational methods for determining dimensionality in exploratory factor analysis. The evaluation was based on each method's ability to indicate the correct number of factors in population correlation matrices with known factor structures, using samples generated through Monte Carlo simulation. In the section that follows, the results reported in the previous chapter will be related back to each of the six research questions outlined toward the end of Chapter One. Then, the results will be articulated within the framework of each specific computational procedure under examination in the present work. A discussion of the implications of these findings, as well as relevant recommendations to factor analysts, limitations of the study design, and suggestions for future research will round out this investigation into the number of factors problem.

#### Resolution of Research Questions

In Chapter One, six empirically testable research questions were developed to guide the current investigation of potential solutions to the number of factors problem. Now, given the observations reported in Chapter Four, each will be systematically addressed.

- Q1 Do any of the three leading computational procedures have difficulty identifying the dimensionality of population correlation matrices with known, simple factor structures?

A desirable quality for any computational method for solving the number of factors problem would be that it have little trouble correctly identifying the true number of factors present in population correlation matrices. In the present study, only PA was able to correctly identify the number of factors present in all 1188 population matrices generated. The accuracy of each of the other methods depended upon whether the eigenvalues being analyzed were calculated from a reduced or from an unreduced correlation matrix.

- Q2 Which of the three procedures most accurately recovers dimensionality using sample correlation matrices drawn from populations with known, simple factor structures?

In the present head-to-head comparison of three computational procedures for determining the number of factors to extract in exploratory factor analysis, PA was observed to most accurately recover dimensionality. This generalization was uniformly true over a wide variety of typical psychometric situations. The differences between the two types of PA studied here, and some of the observed limitations of this procedure will be discussed in more detail in the next section.

Q3 How does the eigenvalues-greater-than-one rule perform in comparison to the other procedures?

The first step of the Little Jiffy approach (Kaiser, 1970) involves retaining all principal components with eigenvalues greater than unity. Prior research has shown that this solution to the number of factors problem is inaccurate when conducting an analysis of sample correlation matrices, and usually results in the erroneous over-extraction of factors (Fava & Velicer, 1992a; Revelle & Rocklin, 1979; Zwick & Velicer, 1982, 1986). Not surprisingly, the present work found that Kaiser's approach performed just as the literature anticipated; it was generally quite inaccurate, with a tendency toward significant over-extraction. In contrast, using the eigenvalues-greater-than-one rule on characteristic roots computed from a reduced sample correlation matrix was found to be a more defensible approach. This variant was found to be fairly accurate, resulting in some slight under-extraction when missing the mark. In general, however, the accuracy of the eigenvalues-greater-than-one rule paled in comparison to the accuracy of either incarnation of PA, both of which towered above all other procedures.

Q4 How does the addition of unique (noise) indicator variables affect the performance of the methods?

The potential presence of unique indicator variables is common in applied psychometric research. A desirable characteristic of any candidate method for solving the numbers of factors problem would be a demonstrated resiliency against the presence of indicator variables which are uncorrelated with any common factors. Of all the various

methods tested in the present study, only the principal components-based variation of PA demonstrated this ability, while maintaining a high level of accuracy.

Q5 Does any single procedure perform well over a wide variety of typical psychometric situations?

Clearly, the principal components version of PA is the superior method for determining the number of factors present in a target population matrix across the wide variety of psychometric situations simulated in this study. Its accuracy was found to be exceptionally high, and the method was even found to be resilient to the presence of unique indicator variables. The one weakness with the method appears to be in the presence of more than eight highly correlated factors, a situation that was also seen to greatly degrade the performance of the other methods.

Q6 What effect does using the reduced correlation matrix (i.e., applying a common factor model) versus using the unreduced correlation matrix (principal components) have on the accuracy of the methods?

The type of correlation matrix used when calculating the eigenvalues used by the computational methods for determining the number of factors had a major effect on the accuracy of the methods. Different methods were affected in different ways. As argued in previous chapters, the method used to determine the number of factors to extract should be the one empirically demonstrated to be the most accurate across a wide range of psychometric situations, and can be employed independently of the model ultimately applied in the factor analysis. For example, a viable approach would be to use the principal components-based version of PA as a “first-pass” to accurately determine the

number of factors for subsequent extraction in another type of common factor analysis—perhaps one based upon maximum likelihood calculations.

### Performance of the Procedures

#### *Standard Error Scree Test*

The standard error scree test, as originally described by Zoski and Jurs (1996), was later evaluated by Nasser et al. (2002) against other regression-based variations of Cattell's (1966) visual scree test. In Nasser et al., the promising SEscree test emerged as the most accurate of the regression-based methods, and its superior performance was found to warrant further investigation alongside more rigorously examined methods, such as PA and MAP. One purpose of the present study was to implement a head-to-head comparison of the three.

Using SAS/IML code adapted directly from Nasser et al. (2002), replication was indeed achieved across the limited combination of variables and conditions outlined in the original study. However, in an attempt to extend those results across a wider variety of typical psychometric situations, the performance of the SEscree was observed to severely deteriorate. This irregularity was detected not only as the number of factors increased, but also as the number of items per factor increased. One possible source of this marked performance decline for SEscree may involve the total number of variables (items), which is a function of both the number of factors and the number of items per factor. For reference, the maximum number of total items examined in Nasser et al. was only 48. Here, conditions were scrutinized where the total number of items exceeded 100.

The results reported above may expose a weakness in the choice of arbiter ( $1/m$ ) recommended by Zoski and Jurs (1996). Those authors assert that an arbiter should be



inversely related to the number of characteristic roots of the correlation matrix under examination, citing an observation by Horn and Engstrom (1979) in which the proportion of error variance tended to be inversely related to sample size. The arbiter described by Zoski and Jurs, then, is a function of the total number of items in the correlation matrix. However, according to empirical evidence generated by this study, when this total number of items is large, the accuracy of the SEscree procedure was observed to decline, and the procedure began to recommend over-factoring—often to an extreme. Perhaps an adjustment of the arbiter is necessary, or perhaps the assumption of the inverse relationship between lack-of-fit error variance and  $m$  is simply erroneous. The investigation of such questions, being thus raised, will not be considered here. For the purposes of the present work, it is sufficient to conclude that the SEscree procedure, as implemented by Zoski and Jurs, and as advocated by Nasser et al. (2002), can be eliminated as a practicable solution to the number of factors problem in typical psychometric research.

#### *Minimum Average Partial Test*

The principal components-based version of Velicer's (1976) MAP test was found to perform reasonably well when applied to sample correlation matrices, in agreement with prior research. Unfortunately, MAP was observed to have trouble identifying the true dimensionality of population matrices. This phenomenon may be viewed as an intuitively unappealing characteristic of any purported solution to the number of factors problem. Additionally, the accuracy of principal components-based MAP was never observed to exceed the accuracy of principal components-based PA.

The common factors-based analogue of MAP, while performing flawlessly for population matrices, flatly blew up when evaluating sample correlation matrices. Recall that this experimental procedure utilized residuals calculated from a principal axis common factoring of each correlation matrix, based on recommendations from Gorsuch (personal communication, February 18, 2008). Closer scrutiny into this method reveals that no residuals can be calculated for a factor whose eigenvalues are less than zero. Negative eigenvalues are common when SMCs are on the diagonal of the correlation matrix, because the matrix is no longer of full rank. Originally, it was thought that this situation would not negatively impact the efficacy of the common factors-based version of MAP; however, given its poor empirical performance, the negative eigenvalues situation must be considered as one conceivable cause. Development and troubleshooting of unproven computational methods for solving the number of factors problem was outside the scope of the current investigation, however, and it is satisfactory to report simply that this particular incarnation of a common factors-based MAP was not the most accurate procedure evaluated in the present study.

### *Parallel Analysis*

The most accurate procedure evaluated in the present study was the principal components-based version of Horn's (1965) parallel analysis. This public domain SAS/IML application—developed by O'Connor (2000a)—was clearly the most accurate computational procedure for solving the number of factors problem under evaluation here. An appealing characteristic of this procedure was its ability to identify the correct number of factors present in the target population matrices without fail. The one weakness of this method surfaced in the psychometrically undesirable situation where

salient factor loadings were low, the number of factors in the target population was high, and the interfactor correlations were oblique. Conceivably, if such conditions were suspected a priori, PA could still prove valuable—as long as the factor analyst employed PA with the caveat that it will tend to recommend under-factoring in this situation. The factor analyst would at least gain a substantive idea of the lower limit for the true number of factors, and be provided with a reasonable starting point for a cogent investigation of dimensionality.

The common factors-based adaptation of PA was also found to be highly accurate. Like the principal components-based construction, it was flawlessly accurate when evaluating population matrices. When considering sample matrices drawn from the population targets, the common factors-based incarnation proved only slightly less accurate than the principal components-based version over a wide variety of psychometric situations. This empirical evidence contradicts predictions made by Crawford and Koopman (1973).

#### Study Limitations and Suggestions for

#### Future Research

Although one intention of this study was to construct target population matrices to represent a wide variety of applied psychometric situations and conditions, the true complexity of most factor analytic research extends far beyond the scope of any one Monte Carlo study. Assumptions and conventions present in the design of this dissertation include normality in the indicator variables, simple factor structure, and equal numbers of items per factor. One cannot assume that PA will perform well outside of these constraints. For example, a major limitation of the present design was an

assumption of continuity in the measured items. Quite often in applied psychometric studies, discrete measures (such as items anchored to a Likert scale) or dichotomous variables (true-false questions) are used instead. Research extending the results of this study to other such typical situations is still necessary.

Turner (1998) and Beauducel (2001) caution that PA may underestimate the number of factors or components to retain when the first eigenvalue is large. Future studies should focus on this potential weakness by manipulating the number of indicator items per factor to vary within each population correlation matrix. Also, the distribution of the factor loadings, although reflecting varied saturation in the present design, could have been varied in different ways. For example, a future design might involve varying the item saturations randomly within the high, wide, and low limits, instead of uniformly dividing these ranges. In the present work, the number of unique indicator variables added was set equal to the number of items per factor—another possible design weakness. The consequences of increased complexity in the factor patterns should also be investigated. For example, how does PA perform when the data contain indicator variables that load on more than one factor? One observation—described by Beauducel, and replicated in the present work—implies that PA may begin to underestimate the true number of factors when dealing with data sets that have oblique simple structure and more than eight factors. Another avenue of future study would be to reexamine the performance of PA, focusing on situations when more than eight or ten factors are present in the target population.

As mentioned previously, the indicator variables generated for this study were standardized to follow a normal distribution. Such is rarely the case for indicator

variables in practice. An investigation into the performance of PA in the presence of non-normal indicator variables is surely warranted. The current design employed the parametric construction of PA, which Buja and Eyuboglu (1992) indicate to be robust against departures from normality. However, those authors also describe a non-parametric variation of PA, which O'Connor (2000a) has conveniently developed alongside the parametric version—it is actually available via an option switch in the same SAS/IML code. Research into the accuracy of non-parametric PA is another necessary direction for future empirical exploration.

Another potential limitation of the present work involves the convention of using eigenvalues from a principal axis factoring to represent common factors eigenvalues. Certainly, numerous alternate common factors procedures exist, which may have produced different sets of eigenvalues—potentially with contrasting results. Investigations into the use of eigenvalues other than those resultant from a principal axis factoring are warranted. Future research could focus on using eigenvalues computed using any factor analytic procedure which does not require the number of factors to be specified a priori.

#### Implications and Recommendations

In the present study, the principal components-based version of PA was found to be the superior solution to the classic number of factors problem. This method had an overall accuracy of over 90 percent, and was within one factor of the correct solution most of the time. Some weaknesses were manifest when the number of factors was high, the interfactor correlations were high, and the range of salient factor loadings was low.

However, when confronted with this situation, PA was observed to only slightly under-factor—by an average of around three to four factors.

Based on the empirical evidence presented here, it is suggested that a factor analyst can confidently adopt the use of PA (based on principal components eigenvalues) as a first-pass solution to the number of factors problem in many commonly encountered psychometric applications of exploratory factor analysis. Upon reaching a solution using PA, the factor analyst can revise this estimate according to previously existing theory, situational considerations, and/or any other traditionally available criteria. Then, the factor analyst is well-prepared to proceed with a second-stage analysis, using the factor analytic procedure of his or her choice.

The continued use of the SEscree procedure is not supported by evidence presented here, especially when the number of items under consideration is high. The use of the principal components version of MAP may be employed as an adjunct, with the caveat that PA is generally much more accurate. Use of the Kaiser (1960) criterion, especially as defined by the Little Jiffy heuristic, is also not supported by the evidence reported here. However, this rule, when used in conjunction with common factors-based eigenvalues, actually performs better than first anticipated. In fact, it rivals the performance of the common factors-based version of PA in certain situations, such as when the number of items per factor is high, and when unique indicator variables are present. However, the performance of this variant of the Kaiser rule was calculated only for completeness here, as its actual use is considered theoretically unsound (see Fabrigar et al., 1999, p. 278; see also Gorsuch, 1980; Gorsuch, 2003; Guttman, 1954; Horn, 1969). Even so, the positive observed results for the common factors-based version of the

eigenvalues-greater-than-one rule may indicate potential utility in investigations of dimensionality; its use in certain specific situations is now at least empirically defensible. Further scrutiny into this matter is necessary.

The results of the present study offer yet another empirical demonstration of the effectiveness and accuracy of the powerful PA procedure. Although not infallible in its present form, one would think that a preponderance of the empirical evidence would at some point be established such that the use of PA would begin to proliferate more in practice. Realistically, however, this proliferation may be slow to manifest until such time as PA becomes a more-easily-implemented proprietary option in the major statistical software packages. As the evidentiary support mounts for the use of parallel analysis as the best first-pass solution to the number of factors problem, it is hereby propounded that the time for action is close at hand.

## REFERENCES

- Allen, S. J., & Hubbard, R. (1986). Regression equations for the latent roots of random data correlation matrices with unities on the diagonal. *Multivariate Behavioral Research, 21*, 393-398.
- American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin, 51*(2, Pt. 2), 1-38.
- Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 111-150.
- Armstrong, J. S., & Soelberg, P. (1968). On the interpretation of factor analysis. *Psychological Bulletin, 70*(5), 361-364.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, 3*, 77-85.
- Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology, 4*, 1-2.
- Beauducel (2001). Problems with parallel analysis in data sets with oblique simple structure. *Methods of Psychological Research Online, 6*(2), 141-157.
- Benson, J., & Nasser, F. M. (1998). On the use of factor analysis as a research tool. *Journal of Vocational Education Research, 23*(1), 13-33.



- Bentler, P. M. (1989). *EQS Structural equations program manual*. Los Angeles: BMDP Statistical Software.
- Bentler, P. M., & Kano, Y. (1990). On the equivalence of factors and components. *Multivariate Behavioral Research*, 25(1), 67-74.
- Borgatta, E. F. (1989). A note on using factor analysis and other procedures in research: A reply to Wilkinson. *Sociological Methods and Research*, 17(4), 460-464.
- Borgatta, E. F., Kercher, K., & Stull, D. E. (1986). A cautionary note on the use of principal components analysis. *Sociological Methods and Research*, 15(1-2), 160-168.
- Box, G. E. P., & Muller, M. E. (1958). A Note on the generation of random normal deviates. *Annals of Mathematical Statistics*, 29(2), 610-611.
- Browne, M. W. (1968). A comparison of factor analytic techniques. *Psychometrika*, 33, 267-334.
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509-540.
- Castellan, J. N. Jr. (1992). Shuffling arrays: Appearances may be deceiving. *Behavior Research Methods, Instruments, and Computers*, 24(1), 72-77.
- Cattell, R. B. (1953). A quantitative analysis of the changes in culture patterns of Great Britain, 1837-1937, by *P* technique. *Acta Psychologica*, 9, 99-121.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

- Cattell, R. B., & Adelson, M. (1951). The dimensions of social change in the U.S.A. as determined by *P* technique. *Social Forces*, 30, 190-201.
- Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*, 12(3), 289-325.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3) 309-319.
- Cliff, N. (1970). The relation between sample size and population characteristic vectors. *Psychometrika*, 35, 163-178.
- Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18(1), 115-126.
- Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin*, 103(2), 276-279.
- Cliff, N., & Hamburger, C. D. (1967). The study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin*, 68(6), 430-445.
- Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46(4), 648-659.
- Comrey, A. L., & Lee, H. B. (1992). A first course in factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147-168.

- Crawford, C. B., & Koopman, P. (1973). A note on Horn's test for the number of factors in factor analysis. *Multivariate Behavioral Research*, 8(1), 117-125.
- Crawford, C. B., & Koopman, P. (1979). Inter-rater reliability of scree test and mean square ratio test of number of factors. *Perceptual and Motor Skills*, 49, 223-226.
- Cronbach, L. J. (1971). Test Validation. In R. L. Thorndike, (Ed.). *Educational measurement* (2nd ed.) (pp. 443-507). Washington, D.C.: American Council on Education.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- DeVellis, R. F. (2003). *Scale development: Theory and application* (2nd ed.). Thousand Oaks, CA: Sage.
- 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. (1992a). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research*, 27(3), 387-415.
- Fava, J. L., & Velicer, W. F. (1992b). An empirical comparison of factor, image, component, and scale scores. *Multivariate Behavioral Research*, 27(3), 301-322.
- Fava, J. L., & Velicer, W. F. (1996). The effects of underextraction in factor and component analysis. *Education and Psychological Measurement*, 56(6), 907-929.
- Finch, J. F., & West, S. G. (1997). The investigation of personality structure: Statistical models. *Journal of Research in Personality*, 31(4), 439-485.

- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291-314.
- Garson, D. G. (2007). *Factor analysis*. Retrieved from:  
<http://www2.chass.ncsu.edu/garson/pA765/factor.htm>
- Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling*, 3, 62-72.
- Glass, G. V., & Taylor, P. A. (1966). Factor analytic methodology. *Review of Educational Research*, 36(5), 566-587.
- Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Education and Psychological Measurement*, 55(3), 377-393.
- Goodwin, L. D., & Goodwin, W. L. (1999). Measurement myths and misconceptions. *School Psychology Quarterly*, 14, 408-427.
- Gorsuch, R. L. (1973). Using Bartlett's significance test to determine the number of factors to extract. *Educational and Psychological Measurement*, 33, 361-364.
- Gorsuch, R. L. (1980). Factor score reliabilities and domain validities. *Educational and Psychological Measurement*, 40(4), 895-894.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, N.J.: Lawrence Earlbaum.

- Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, 25(1), 33-39.
- Gorsuch, R. L. (1997). 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, (Vol. Eds.). *Handbook of psychology. Vol. 2: Research methods in psychology*. (pp. 143-164). Hoboken, N. J.: John Wiley & Sons.
- Gorsuch, R. L., & Nelson, J. (1981). CNG scree test: An objective procedure for determining the number of factors. Paper presented at the annual meeting of the Society for Multivariate Experimental Psychology.
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-275.
- Guilford, J. P. (1946). New standards for test evaluation. *Educational and Psychological Measurement*, 6(4), 427-439.
- Guttman, L. (1940). Multiple rectilinear prediction and the resolution into components. *Psychometrika*, 5, 75-99.
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19, 149-161.
- Guttman, L. (1955). The determinacy of factor score matrices with implications for five basic problems of common-factor theory. *British Journal of Statistical Psychology*, 8, 65-81.
- Hakstian, A. R., & Muller, V. J. (1973). Some notes on the number of factors problem. *Multivariate Behavioral Research*, 8(4), 461-475.

- Harris, M. L., & Harris, C. W. (1971). A factor analytic interpretation strategy. *Educational and Psychological Measurement, 31*, 589-606.
- 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.
- Hong, S. (1999). Generating correlation matrices with model-error for simulation studies in factor analysis: A combination of the Tucker-Koopman-Linn model and Wijsnman's algorithm. *Behavior Research Methods, Instruments, and Computers, 31*(4), 727-730.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185.
- Horn, J. L. (1969). On the internal consistency reliability of factors. *Multivariate Behavioral Research, 4*(1), 115-125.
- Horn, J. L., & Engstrom, R. (1979). Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research, 14*(3), 283-300.
- Hubbard, R., & Allen, S. J. (1987a). An empirical comparison of alternative methods for principal component extraction. *Journal of Business Research, 15*(2), 173-190.
- Hubbard, R., & Allen, S. J. (1987b). A cautionary note on the use of principal components analysis: Supportive empirical evidence. *Sociological Methods and Research, 16*(2), 301-308.

- Hubbard, R., & Allen, S. J. (1989). Differences in empirical results using principal components and factor analysis: A reply to Wilkinson. *Sociological Methods and Research*, 17(4), 465-466.
- Humphreys, L. G., & Illgen, D. R. (1969). Note on a criterion for the number of common factors. *Educational and Psychological Measurement*, 29, 571-578.
- Humphreys, L. G., & Montanelli, R. G. Jr. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, 10, 193-205.
- Humphreys, L. G., Ilgen, D., McGrath, D., & Montanelli, R. (1969). Capitalization on chance in rotation of factors. *Educational and Psychological Measurement*, 29; 259-271.
- Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., et al. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior*, (18)6, 667-683.
- 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.
- Jackson, D. A. (1993). Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches. *Ecology*, 74(8), 2204-2214.
- Jackson, D. N., & Chan, D. W. (1980). Maximum-likelihood estimation in common factor analysis: A cautionary note. *Psychological Bulletin*, 88(2), 502-508.

- Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. Upper Saddle River, N. J.: Prentice-Hall.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183-202.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.
- Kaiser, H. F. (1956). *The varimax method of factor analysis*. Unpublished doctoral dissertation, University of California, Berkeley.
- Kaiser, H. F. (1958). The varimax criteria for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika*, 35(4), 401-415.
- Kaiser, H. F., & Dickman, K. (1962). Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. *Psychometrika*, 27(2), 179-182.
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111-117.
- Kanyongo, G. Y. (2005). Determining the correct number of components to extract from a principal components analysis: A Monte Carlo study of the accuracy of the scree plot. *Journal of Modern Applied Statistical Methods*, 4(1), 120-133.
- Karr, R. J., & Martin, T. E. (1981). Random numbers and principal components: Further searches for the unicorn. In D. E. Capen, (Ed.). *The use of multivariate statistics*



- in studies of wildlife habitat*. (pp. 20-24). United States Forest Service General Technical Report RM-87.
- Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151-160.
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organization Research Methods*, 9(2), 202-220.
- Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh*, 61, 176-185.
- Lawley, D. N. (1941). Further investigation of factor estimation. *Proceedings of the Royal Society of Edinburgh*, 61, 176-185.
- Lee, H. B., & Comrey, A. L. (1979). Distortions in a commonly used factor analytic procedure. *Multivariate Behavioral Research*, 14(3), 301-321.
- Leners, D. W., Roehrs, C., & Piccone, A. V. (2006). Tracking the development of professional values in undergraduate nursing students. *Journal of Nursing Education*, 45(2), 504-511.
- Levonian, E., & Comrey, A. L. (1966). Factorial stability as a function of the number of orthogonally-rotated factors *Behavioral Science*, 11(5), 400-404.
- Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. *Psychometrika*, 33(1), 37-71.

- Longman, R. S., & Holden, R. R. (1992). PERM: A double-precision FORTRAN routine for obtaining parallel analysis eigenvalues from permuted data. *Behavior Research Methods, Instruments, and Computers*, 24(1), 493.
- 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(1), 59-69.
- 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.
- MacCallum, R. C. (1983). A comparison of factor analysis programs in SPSS, BMDP, and SAS. *Psychometrika*, 48, 223-231.
- MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin*, 109(3), 502-511.
- Merenda, P. F. (1997). A guide to the proper use of factor analysis in the conduct and reporting of research: Pitfalls to avoid. *Measurement and Evaluation in Counseling and Development*, 30(3), 156-164.
- Montanelli, R. G. Jr., & Humphreys, L. G. (1976). Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika*, 41(3), 341-348.
- Mundfrom, D., Shaw, D., & Ke, T. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, 5(2), 159-168.

- Nasser, F. M. (1997). *The performance of regression-based variations of the visual scree for determining the number of common factors*. Unpublished doctoral dissertation, University of Georgia, Athens.
- Nasser, F. M., 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(3), 397-419.
- Natsopoulos D., Kiosseoglou, G., Xeromeritou, A., & Alevriadou, A. (1998). Do the hands talk on mind's behalf? Differences in language ability between left- and right-handed children. *Brain and Language*, 64, 182-214.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- O'Connor, B. P. (2000a). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers*, 32(3), 396-402.
- O'Connor, B. P. (2000b). *Using parallel analysis and Velicer's MAP test*. Retrieved February 28, 2009, from <http://flash.lakeheadu.ca/~boconno2/nfactors/nfactors.html>
- Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics and Data Analysis*, 49, 974-997.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.

- Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, 2(1), 13-43.
- Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors *Multivariate Behavioral Research*, 14, 403-414.
- SAS program for determining the number of components using parallel analysis. (2006). [Computer software]. Retrieved from <http://flash.lakeheadu.ca/~boconno2/nfactors/rawpar.sas>
- SAS program for determining the number of components using Velicer's MAP test. (2006). [Computer software]. Retrieved from <http://flash.lakeheadu.ca/~boconno2/nfactors/map.sas>
- Schönemann, P. H. (1990). Facts, fictions, and common sense about factors and components. *Multivariate Behavioral Research*, 25(1), 47-51.
- Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th ed.). Ames, IA: Iowa State University Press.
- Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, 106(1), 148-154.
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, (15)2, 201-292.
- Spearman, C. (1927). *The abilities of man*. New York: MacMillan.
- Steiger, J. H., & Schönemann, P. H. (1978). A history of factor indeterminacy. In S. Shye, (Ed.). *Theory construction and data analysis in the behavioral sciences*. (pp. 136-178). San Francisco: Jossey-Bass.

- Stewart, D. W. (1981). The application and misapplication of factor analysis in marketing research. *Journal of Marketing Research*, 18, 51-62.
- Streiner, D. L. (1998). Factors affecting the reliability of interpretations of scree plots. *Psychological Reports*, 83, 687-694.
- Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, 56(2), 197-208.
- Thurstone, L. L. (1935). *The vectors of the mind*. Chicago: The University of Chicago Press.
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, 34(4), 421-459.
- Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement*, 58(4), 541-568.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321-327.
- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2), 231-251.
- Velicer, W. F., & Jackson, D. N. (1990a). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25(1), 1-28.

- Velicer, W. F., & Jackson, D. N. (1990b). Component analysis versus common factor analysis: Some further observations. *Multivariate Behavioral Research*, 25(1), 97-114.
- 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*. (pp. 41-71). Boston: Kluwer Academic Publishers.
- Velicer, W. F., Peacock, A. C., & Jackson, D. N. (1982). A comparison of component and factor patterns: A Monte Carlo approach. *Multivariate Behavioral Research*, 17(3), 371-388.
- Watkins, M. W. (2006). Determining parallel analysis criteria. *Journal of Applied Statistical Methods*, 5(2), 344-346.
- Weiss, D. (1976). Multivariate Procedures. In M. D. Dunnette (Ed.). *Handbook of Industrial and Organizational Psychology*. (pp. 327-362). Chicago: Rand McNally.
- Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28(3), 263-311.
- Wilkinson, L. (1986). *SYSTAT*. Evanston, IL: Systat.
- Wilkinson, L. (1989a). A cautionary note on the use of factor analysis: A response to Borgatta, Kercher, and Stull, and Hubbard and Allen. *Sociological Methods and Research*, 17(4), 449-459.

- Wilkinson, L. (1989b). A final note on Borgatta, Kercher, and Stull, and Hubbard and Allen. *Sociological Methods and Research*, 17(4), 467.
- 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.
- Yeomans, K. A., & Golder, P. A. (1982). The Guttman-Kaiser criterion as a predictor of the number of common factors. *The Statistician*, (31)3, 221-229.
- Zoski, K. W., & Jurs, S. (1993). Using multiple regression to determine the number of factors to retain in factor analysis. *Multiple Linear Regression Viewpoint*, 20, 5-9.
- Zoski, K. W., & 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(2), 253-269.
- 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.

## APPENDIX A

### ANNOTATED SAS/IML CODE



```

DM LOG 'CLEAR';                                / *Clear Log Window */
DM OUTPUT 'CLEAR';                             / *Clear Output Window */

PROC DATASETS LIB=WORK KILL NOLIST; RUN; QUIT; /* Kill Data & Graphs */
GOPTIONS RESET=GLOBAL;                        /* Reset Graph Options */

OPTION nonotes nomprint NODATE PAGENO = 1 nosource;

%let path=T:\Adam\ml low\;

libname SIM "&path";
filename dummyout dummy 'any_file_name';

proc printto log="&path.GENCORR.log" new print=dummyout;
run;

%macro GENCORR(f,nv,p,r,m,u,n,s,np,per,ty,sh);

title1;
title2;

data time; /* Initialize start time for each condition */
time1 =datetime();
run;

/*****

%if &u eq 1 %then %let g = &f; /* No Unique Factor */
%if &u eq 2 %then %let g = %eval(&f+1); /* Add a Unique Factor */

proc iml;
reset fuzz; /* Rounds extremely small values to zero */

/* Input pattern matrix*/
do i=1 to &g; /* Creates and concatenates g factor matrices into a
pattern matrix */

    x=j(&nv, &g, 0); /* Creates nv x g submatrix of zeros */
    y=j(&nv, 1, &p); /* Creates nv x 1 column vector of p's */

    min_loading=&m; /* Minimum loading for varied saturation
condition */
    max_loading=&p; /* Maximum loading for varied saturation
condition */
    interval = (max_loading-min_loading)/(nrow(y)-1);

    do j=1 to &nv;

        y[j,1]=round(max_loading-((j-1)*interval),.01); /* Loadings for
varied saturation condition, rounded to hundredths */

```

```

        if (&u = 2 & &g = i) then y[j,1]=0; /* Unique Factor Loadings are
Zero */

        end;

        x[,i]=y; /* Inserts p's into the ith column of the nv x g
submatrix */
        p=p//x; /* Vertically concatenates each factor submatrix into the
pattern matrix, p */
        u=u//y; /* Vertically concatenates each loading into a vector of
for computing uniquenesses, u */

end;

if (&u = 2) then p=p[,1:ncol(p)-1]; /* If Adding a Unique Factor, Trim
Extra Column */

v=nrow(p); ** Total number of variables;
call symput('v',strip(char(v)));

* need to output pattern matrix;

*PRINT 'Pattern Matrix', p ; /* Pattern matrix */

x=j(ncol(p),ncol(p),&r); /* Creates square matrix of r's */
c=x<>i(ncol(p)); /* Creates interfactor correlation matrix, c */

*print 'Interfactor Correlation Matrix', c; /* Interfactor Correlation
Matrix */

u=diag(1-u##2); /* Compute Diagonal Matrix of Uniquenesses */

*PRINT 'Diagonal Matrix of Uniquenesses', u ; /* Diagonal Matrix of
Uniquenesses */

r=p * c * T(p) + u; /* Adds a matrix of SMC's as communalities,
guarantees 1's in all diagonals */

*PRINT 'Population Correlation Matrix', r ; /* Population Correlation
Matrix */

create temp from r; /* Creates SAS dataset */
append from r;

quit;

data popcorr; /* Population correlation matrix */
set temp;
length _type_ $ 8;
_type_ = 'CORR';
run;

data popcorr; /* Adds name variable to population correlation matrices
*/

```

```

set popcorr;
length _name_ $ 8;
do i=1 to &v;
  _name_=strip('COL'||put(_n_, 8.-L));
  call symput('name'||put(_n_, 8.-L),_name_);
end;
drop i;
run;

data descriptives1; /* Calculates descriptives to population
correlation matrix */
length _type_ $ 8;
_type_='MEAN';
%do nameindex=1 %to &v;
  &&name&nameindex=0;
%end;
output;
_type_='STD';
%do nameindex=1 %to &v;
  &&name&nameindex=1;
%end;
output;
_type_='N';
%do nameindex=1 %to &v;
  &&name&nameindex=10000;
%end;
output;
run;

data popcorr(type=corr); /* Reorder variables and specify dataset as a
correlation matrix */
retain _type_ _name_;
set descriptives1 popcorr;
run;

*proc print data=popcorr; /* Population correlation matrix */

/* Create principal components matrix from which sample will be
generated (Kaiser-Dickman Method)*/
proc factor data=popcorr n=&v outstat=patmat noprint;
var coll-col&v;
run;

*proc print data=patmat;

data patmat;
set patmat;
if _type_='PATTERN';
run;

*proc print data=patmat; /* Population Principal Components Matrix */

proc iml;
reset fuzz;
seed=&seed; /* Random number seed */

/* Create SAS datasets */

```

```

dum = j(1,&v+1,0);
create rout1 from dum;
create rout2 from dum;

/* Import data and create population pattern matrix, f */
use patmat;
read all into f;

*print 'Principal Components Matrix', f; /* Principal Components Matrix
*/

/* Generate sample correlation matrices (n samples of size s)*/
do loop=1 to &n;
    lvec1=loop*j(&v,1,1); /* Creates a v x 1 column vector of index
numbers */
    lvec2=loop*j(&s,1,1); /* Creates an s x 1 column vector of index
numbers */

    d= normal(j(&s,&v,seed)) * f; /* Generates an s x v matrix of
pseudo-random standard normal numbers
                                (based on the Box-Muller
transformation of the UNIFORM function)
                                and then multiplies it by the
population pattern matrix to create
                                a random sample data matrix, d
(Kaiser-Dickman Method)*/

    *print d; /* Sample data matrix */

    nn=nrow(d); /* Finds sample size */
    sum=d[+,]; /* Creates a 1 x s column vector of column sums */
    xpx= (T(d)*d - T(sum)*sum/nn)/(nn-1);
    dg=inv(sqrt(diag(xpx)));
    r=dg*xpx*dg; /* Computes sample correlation matrix, r */

    dum1=r || lvec1; /* Concatenate an index vector to each
correlation matrix */
    setout rout1; /* Sets rout1 for data output */
    append from dum1; /* Append each correlation matrix to a SAS
data set */

    dum2=d || lvec2; /* Concatenate an index vector to each sample
data matrix */
    setout rout2; /* Sets rout2 for data output */
    append from dum2; /* Append each sample data matrix to a SAS data
set */

end;

store seed; /* Retain seed, free all other variables */
free / seed;

quit;

```

```

data work.corr;          /* Saves data set, n vertically concatenated
correlation matrices with index column */
set rout1;
length _type_ $ 8;
_type_='CORR';
index=4;
run;

data work.raw;          /* Saves data set, n vertically concatenated
sample data matrices with index column */
set rout2;
run;

data corr;  /* Adds name variable to correlation matrices */
set corr;
length _name_ $ 8;
offset=(col%eval(&v+1)-1)*&v;
_name_=left(trim('COL'||put(_n_-offset, 8.-L)));
drop offset;
run;

data corr;  /* Reorder variables */
retain _type_ _name_ ;
set corr;
run;

proc means data=work.raw noprint; /* Calculates descriptives for
correlation matrices */
by col%eval(&v+1);
output out=descriptives2;
run;

data descriptives2;          /* Selects appropriate descriptives
for data set */
set descriptives2;
if _stat_='MIN' then delete;
if _stat_='MAX' then delete;
drop _type_ _freq_;
run;

data descriptives2; /* Rename variables correctly */
set descriptives2;
_type_=_stat_;
if _type_='MEAN' then index=1;
if _type_='STD' then index=2;
if _type_='N' then index=3;
drop _stat_;
run;

proc sort data=descriptives2; /* Sort */
by col%eval(&v+1);
run;

data corr; /* Add descriptives to data set */
set corr descriptives2;
by col%eval(&v+1);
run;

```

```

proc sort data=corr; /* Sort descriptives */
by col%eval(&v+1) index;
run;

data temp2;
set popcorr;
COL%eval(&v+1)=0;
run;

data corr(type=corr); /* Specify as a correlation matrix */
set temp2 corr;
drop index;
run;

/***** End Data Generation *****/

/***** Start SESCOree *****/

%let priors=;

/* Principal Components Analysis of the n sample correlation matrices */
%if &ty eq 1 %then %do;
%let priors=one;
%end;

/* Principal Axis Factor Analyze the n sample correlation matrices */
%if &ty eq 2 %then %do;
%let priors=smc;
%end;

/* Compute eigenvalues for population and all sample correlation
matrices */
proc factor data=work.corr method=principal n=&v rotate=varimax
priors=&priors outstat=eigen noprint riter=10000;
var coll-col&v;
by col%eval(&v+1); /* By variable is the index column */
run;

data eigen;
set eigen;
keep coll-col&v;
if _type_='EIGENVAL';
run;

*** Experimental MAP Part ***;

***** No Outputs *****;
ODS RESULTS OFF;
*filename dummyout dummy 'any_file_name';
*proc printto print=dummyout; * Prints results to dummy file;
*run;
*****;

%do resid=1 %to &v;

```

```

proc factor data=work.corr method=principal mineigen=0 n=&resid
priors=&priors residuals; /* Calculate residuals for PAF MAP */
var coll=col&v;
by col%eval(&v+1); /* By variable is the index column */
ods select ParCorrControlFactor;
ods output ParCorrControlFactor = resid&resid;
run;

data resid&resid;
retain extracted;
set resid&resid;
extracted=&resid;
run;

%end;

***** Restore Outputs *****;
ODS RESULTS ON;
*proc printto; * Releases Print file;
*run;
*****;

data resid; /* Initialize */
run;

%do resid=1 %to &v; /* Concatenate residuals */
data resid;
set resid resid&resid;
if coll=. then delete;
run;

proc datasets nolist; /* Clean up */
delete resid&resid;
run;

%end;
*****;

*proc print data=kgeigen; /* Data set of n+1 rows of principal
components eigenvalues for KG benchmark */

*proc print data=eigen; /* Data set of n+1 rows of eigenvalues of the
sample correlation matrices, the first row contains population
eigenvalues */

proc datasets nolist; /* Clean up */
delete temp temp2 popeigen descriptives1 descriptives2;
run;

proc iml;

use eigen;
read all into g; /* Creates an n+1 x v matrix, g, with n row vectors of
eigenvalues of the sample correlation matrices, the first row contains
population eigenvalues */

```

```

*print g; /* Matrix of sample and population eigenvalues */

/***** Start SEscree
*****/

do ij=1 to nrow(g); /* Loop to do procedure for all samples */

    e=t(g[ij,]); /* Creates a v x 1 column vector of eigenvalues, e,
from g */

    *print e; /* Vector of sample eigenvalues */

    f=j(&v,1,1);
    o=j(&v,1,1); /* Creates a v x 1 column vector of 1's */

    do i=1 to &v; /* Creates a v x 1 column vector of ordered
integers from 1 to v */
        f[i]=i;
    end;
    *print f o;

    mrmse=j(%eval(&v-2),1,0); /* Vector of the mean square for scree
*/
    rmse=j(%eval(&v-2),1,0); /* Vector of the mean square residual
(scree) */
    do i=1 to %eval(&v-2);
        y=e[i:&v,];

        x=f[i:&v,]||o[i:&v,];
        *print x;
        b=inv(x`*x)*x`*y; /* 2 x 1 Column vector of betahats */

        print b;

        mp=%eval(&v+1)-i; /* N, Number of variables used in the SE
calculation */
        rmse[i]=(y`*y - b`*x`*y); /* (Y - Yhat)squared */
        if rmse[i]<0 then rmse[i]=0; /* Keep numerator non-
negative */
        mrmse[i]=rmse[i]/(mp-2); /* Denominator is N - Rank of X
matrix (2)*/
        *print rmse mp mrmse;

    end;

    vmrmse=mrmse[><,]; /* Finds element minimum */

    *print vmrmse;

    do i=1 to %eval(&v-2);
        if abs(mrmse[i]-vmrmse)<.000001 then mrmse[i]=vmrmse;
    end;

    sqmrmse=sqrt(mrmse); /* Creates a v x 1 column vector of
standard error of estimates */
    *print ij sqmrmse;

```



```

        fscree=.;

        do i=1 to %eval(&v-2);
            if sqmrmse[i] < (1/&v) then do; /* This is the heart of
SEscree */
                fscree=i-1;
                goto jump1;
            end;

        end;
        jump1: ;

        dum1=dum1//fscree; /* Saves results */

end;

/***** End SEscree *****/

/***** Start MAP *****/

/* Velicer's Minimum Average Partial (MAP) Test */

do i=0 to &n; /* Do all n matrices */

    USE work.corr;
    read all into cr where (col%eval(&v+1)=i);

    cr=cr[4:nrow(cr),1:&v]; /* Extracts v x v sample correlation
matrix, Trims descriptives */

    *call eigen (eigval,eigvect,cr);
    *loadings = eigvect * sqrt(diag(eigval)); /* Can't take square
root of negative eigenvalues, MAP crashes for PAF */

    fm = j(nrow(cr),2,-9999);
    fm[1,2] = (ssq(cr) - ncol(cr))/(ncol(cr)*(ncol(cr)-1));

    do m = 1 to ncol(cr) - 1;
        USE work.resid;
        read all into partials where (col%eval(&v+1)=i &
extracted=m);
        pr=partials[1:nrow(partials),3:ncol(partials)]; /* Extracts
v x v partial correlation matrix, Trims indices */
        fm[m+1,2] = (ssq(pr)-ncol(cr)) / (ncol(cr)*(ncol(cr)-1));
    end;

    *print fm;
    /* identifying the smallest fm value & its location (= the of
factors) */
    minfm = fm[1,2];
    nfactors = 0;
    do s = 1 to nrow(fm);
        fm[s,1] = s - 1;
        if ( fm[s,2] < minfm ) then do;
            minfm = fm[s,2];
            nfactors = s - 1;
        end;
    end;
end;

```

```

end;
end;

/* print, "Velicer's Minimum Average Partial (MAP) Test:";
   print  "Velicer's Average Squared Correlations", fm[format=12.6];
   print  "The smallest average squared correlation is",
minfm[format=12.6];
   print  "The number of components is", nfactors; */

dum2=dum2//nfactors; /* Saves results */

end;

free / dum1 dum2; /* Free all variables except results */

/***** End MAP *****/

/***** Start PA *****/

/* Parallel Analysis Program For Raw Data and Data Permutations

This program conducts parallel analyses on data files in which
the rows of the data matrix are cases/individuals and the
columns are variables; Data are read/entered into the program
using the READ command (see the READ command below);
Alternative procedures for entering data in PROC IML include
the USE, READ, INFILE, INPUT, and EDIT commands;
There can be no missing values;

You must also specify:
-- the # of parallel data sets for the analyses;
-- the desired percentile of the distribution and random
   data eigenvalues;
-- whether principal components analyses or principal axis/common
   factor analysis are to be conducted, and
-- whether normally distributed random data generation or
   permutations of the raw data set are to be used in the
   parallel analyses;

WARNING: Permutations of the raw data set are time consuming;
Each parallel data set is based on column-wise random shufflings
of the values in the raw data matrix using Castellán's (1992,
BRMIC, 24, 72-77) algorithm; The distributions of the original
raw variables are exactly preserved in the shuffled versions used
in the parallel analyses; Permutations of the raw data set are
thus highly accurate and most relevant, especially in cases where
the raw data are not normally distributed or when they do not meet
the assumption of multivariate normality (see Longman & Holden,
1992, BRMIC, 24, 493, for a Fortran version); If you would
like to go this route, it is perhaps best to (1) first run a
normally distributed random data generation parallel analysis to
familiarize yourself with the program and to get a ballpark
reference point for the number of factors/components;
(2) then run a permutations of the raw data parallel analysis
using a small number of datasets (e.g., 10), just to see how long
the program takes to run; then (3) run a permutations of the raw
data parallel analysis using the number of parallel data sets that

```

```

you would like use for your final analyses; 100 datasets are
usually sufficient, although more datasets should be used
if there are close calls. */

/* Define Modules outside of do loop */

/* set diagonal to a column vector module */
start setdiag(matname,vector);
do i = 1 to nrow(matname);
do j = 1 to ncol(matname);
if (i = j) then; matname[i,j] = vector[i,1];
end;end;
finish;

/* row sums module */
start rsum(matname);
rsums =j(nrow(matname),1);
do rows = 1 to nrow(matname);
dumr = matname[rows,];
rsums[rows,1]=sum(dumr);
end;
return(rsums);
finish;

/* Pearson correlation matrix module */
start corrcoef(matname);
ncases = nrow(matname);
nml = 1 / (ncases-1);
vcv = nml * (t(matname)*matname -
((t(matname[+,]) *matname[+,])/ncases));
d = inv(diag(sqrt(vecdiag(vcv))));
r = d * vcv * d;
return(r);
finish;

load seed;

/* Enter your specifications: */

/* Enter or read a raw data matrix, where rows = cases,
& columns = variables
Use the following name for the raw data matrix: "raw".
Cases with missing values are not permitted in the data file. */

do in=0 to &n; /* Do all n matrices */

    if in=0 then do; * Population Data;

        USE work.corr;
        read all into r where (col%eval(&v+1)=0);
        r=r[4:&v+3,1:&v]; /* Imports v x v population
correlation matrix */
        *print r; /* Population correlation matrix */
    end;
end;

```

```

/* Enter the desired number of parallel data sets here */
ndatsets = &np;

/* Enter the desired percentile here */
percent = &per;

/* Specify the desired kind of parellel analysis, where:
   1 = principal components analysis
   2 = principal axis/common factor analysis */
kind = &ty ;

/* Enter either
   1 for normally distributed random data generation
parallel analysis, or
   2 for permutations of the raw data set */
randtype = 1;

/* End of required user specifications */

ncases = 10000;
nvars = ncol(r);

/* principal components analysis & random normal data
generation */
if kind = 1 & randtype = 1 then do;
  realeval = eigval(r);
  evals = j(nvars,ndatsets,-9999);
  do nds = 1 to ndatsets;
    evals[,nds] = eigval(corrcoef(normal(j(ncases,nvars)))));
  end;
end;

/* PAF/common factor analysis & random normal data
generation */
if kind = 2 & randtype = 1 then do;
  smc = 1 - (1 / vecdiag(inv(r)) );
  run setdiag(r,smc);
  realeval = eigval(r);
  evals = j(nvars,ndatsets,-9999);
  do nds = 1 to ndatsets;
    r = corrcoef(normal(j(ncases,nvars))));
    smc = 1 - (1 / vecdiag(inv(r)) );
    run setdiag(r,smc);
    evals[,nds] = eigval(r);
  end;
end;

/* identifying the eigenvalues corresponding to the desired
percentile */
num = round((percent*ndatsets)/100);
results = j(nvars,4,-9999);
results[,1] = t(1:nvars);
results[,2] = realeval;
do root = 1 to nvars;
  ranks = rank(evals[root,]);
  do col = 1 to ndatsets;
    if (ranks[1,col] = num) then do;

```

```

results[root,4] = evals[root,col];
col = ndatsets;
end;
end;
end;
results[,3] = evals[,+] / ndatsets;

reig=results[,2];
rpct=results[,4];

*print reig rpct;

*print realeval; * Eigenvalues from sample data generated
by PA in IML;

do i=1 to nrow(reig);
if reig[i] < rpct[i] then do; /* This is the heart of PA */
    fpa=i-1;
    goto jump2;
end;

end;
jump2: ;

dum3=dum3//fpa; /* Saves results */

end;
else do; * Sample data;

    USE work.raw;
    read all into raw where (col%eval(&v+1)=in);

    raw=raw[1:&s,1:&v]; /* Imports the next s x v sample
data matrix */

    *print raw; /* Raw sample data matrix */

    /* Enter the desired number of parallel data sets here */
    ndatsets = &np;

    /* Enter the desired percentile here */
    percent = &per;

    /* Specify the desired kind of parellel analysis, where:
    1 = principal components analysis
    2 = principal axis/common factor analysis */
    kind = &ty ;

    /* Enter either
    1 for normally distributed random data generation
parallel analysis, or
    2 for permutations of the raw data set */
    randtype = &sh ;

    /* End of required user specifications */

```

```

ncases    = nrow(raw);
nvars     = ncol(raw);

/* principal components analysis & random normal data
generation */
if kind = 1 & randtype = 1 then do;
  realeval = eigval(corrcoef(raw));
  evals = j(nvars,ndatsets,-9999);
  do nds = 1 to ndatsets;
    evals[,nds] = eigval(corrcoef(normal(j(ncases,nvars)))));
  end;
end;

/* principal components analysis & raw data permutation */
if kind = 1 & randtype = 2 then do;
  realeval = eigval(corrcoef(raw));
  evals = j(nvars,ndatsets,-9999);
  do nds = 1 to ndatsets;
    x = raw;
    do lupec = 1 to nvars;
      do luper = 1 to (ncases - 1);
        k = int( (ncases - luper + 1) * uniform(seed) + 1 ) +
luper - 1;
        d = x[luper,lupec];
        x[luper,lupec] = x[k,lupec];
        x[k,lupec] = d;
      end;
    end;
    evals[,nds] = eigval(corrcoef(x));
  end;
end;

/* PAF/common factor analysis & random normal data
generation */
if kind = 2 & randtype = 1 then do;
  r = corrcoef(raw);
  smc = 1 - (1 / vecdiag(inv(r)) );
  run setdiag(r,smc);
  realeval = eigval(r);
  evals = j(nvars,ndatsets,-9999);
  do nds = 1 to ndatsets;
    r = corrcoef(normal(j(ncases,nvars)))));
    smc = 1 - (1 / vecdiag(inv(r)) );
    run setdiag(r,smc);
    evals[,nds] = eigval(r);
  end;
end;

/* PAF/common factor analysis & raw data permutation */
if kind = 2 & randtype = 2 then do;
  r = corrcoef(raw);
  smc = 1 - (1 / vecdiag(inv(r)) );
  run setdiag(r,smc);
  realeval = eigval(r);
  evals = j(nvars,ndatsets,-9999);

```

```

do nds = 1 to ndatsets;
x = raw;
do lupec = 1 to nvars;
do luper = 1 to (ncases -1);
k = int( (ncases - luper + 1) * uniform(seed) + 1 ) +
luper - 1;
d = x[luper,lupec];
x[luper,lupec] = x[k,lupec];
x[k,lupec] = d;
end;
end;
r = corrcoef(x);
smc = 1 - (1 / vecdiag(inv(r)) );
run setdiag(r,smc);
evals[,nds] = eigval(r);
end;
end;

/* identifying the eigenvalues corresponding to the desired
percentile */
num = round((percent*ndatsets)/100);
results = j(nvars,4,-9999);
results[,1] = t(1:nvars);
results[,2] = realeval;
do root = 1 to nvars;
ranks = rank(evals[root,]);
do col = 1 to ndatsets;
if (ranks[1,col] = num) then do;
results[root,4] = evals[root,col];
col = ndatsets;
end;
end;
end;
results[,3] = evals[,+] / ndatsets;

/*
print, "Parallel Analysis:";
if (kind = 1 & randtype = 1) then;
print, "Principal Components & Random Normal Data
Generation";
if (kind = 1 & randtype = 2) then;
print, "Principal Components & Raw Data Permutation";
if (kind = 2 & randtype = 1) then;
print, "PAF/Common Factor Analysis & Random Normal Data
Generation";
if (kind = 2 & randtype = 2) then;
print, "PAF/Common Factor Analysis & Raw Data Permutation";
specifs = (ncases // nvars // ndatsets // percent);
rlabels = {"Ncases" "Nvars" "Ndatsets" "Percent"};
print, "Specifications for this Run:",
specifs[rowname=rlabels];
clabels={"Root" "Raw Data" "Means" "Prctype"};
print "Raw Data Eigenvalues, & Mean & Percentile Random
Data Eigenvalues",
results[colname=clabels format=12.6];
*/

```

```

        reig=results[,2];
        rpct=results[,4];

        *print reig rpct;
        *print realeval; * Eigenvalues from sample data generated
by PA in IML;

        do i=1 to nrow(reig);
        if reig[i] < rpct[i] then do; /* This is the heart of PA */
            fpa=i-1;
            goto jump3;
        end;

        end;
        jump3: ;

        dum3=dum3//fpa; /* Saves results */

    end;
end;

/***** END PA *****/

/***** Start KG Benchmark *****/

use eigen;
read all into kg; /* Creates an n+1 x v matrix, g, with n row vectors
of eigenvalues of the sample correlation matrices, the first row
contains population eigenvalues */

*print kg; /* Matrix of population and sample principal components
eigenvalues */

do i=1 to nrow(kg); /* Loop to do KG procedure for all samples */
    temp=j(ncol(kg),1,0);
    do j=1 to ncol(kg);
        if kg[i,j]>1 then temp[j,1]=1;
    end;
    kgnum=temp[+];
    dum4=dum4//kgnum;
end;

/***** End KG Benchmark *****/

/***** Concatenate and organize results *****/

dum=j(&n+1,1,0);

/* Generates an index for each iteration */
do i=0 to &n;
    dum[i+1]=i;
end;

dum=dum||dum1||dum2||dum3||dum4;

```



```

*print dum;

create results from dum; /* Saves results for all three methods and KG
Benchmark*/
append from dum;

quit;

/***** Time Stamp Code *****/
data time;
set time;

time2=datetime();

time_s=INTCK( 'second', time1, time2 ) ;
time_m=int(time_s/60) ;
time_h=int(time_m/60) ;

time_s=time_s-(time_m*60);
time_m=time_m-(time_h*60);

d1=datepart(time1);
d2=datepart(time2);
t1=timepart(time1);
t2=timepart(time2);

format d1 d2 worddate16.;
format t1 t2 timeampm13.;

put "***** Timer
*****";
put;
put "    Condition &condnum of &numcond ";
put;
put "    Macro started on " d1 "at " t1;
put "    Macro finished on " d2 "at " t2;
put;
put "    Macro took " time_h "Hour(s), " time_m "Minute(s), and " time_s
"Second(s)";
put;
put
"*****"
;

run;
/*****
*/

%mend;

/***** MACRO GENCORR *****/
*
*
*      %GENCORR(f,nv,p,r,m,n,s,np,per,tp,sh)
*

```

```

*
*      f = # of factors
*      nv = # of variables/factor
*      p = Factor loadings
*      r = Interfactor correlations
*      m = Minimum loading
*      u = Unique factors|No Unique factors
*      n = # of sample correlation matrices
*      s = size of each random sample
*
*      np = # of Random Shufflings in PA
*      per = %tile cutoff for PA
*      ty = Principal Components|PAF
*      sh = Random normal|Shuffle raw data
*
*****/

/* (f = 4, v = 24, p = .8, N = 250, r = .4) */

%GENCORR(4, 6, .8, .4, 100, 250, 100); /* Example from Article */
%GENCORR(4, 4, .5, .4, 100, 250, 100); /* Example from Dissertation */

%macro main;

/* Initialize data sets for final output */
data sim.results;
f=0; nv=0; p=0; v=0; r=0; m=0; u=0; n=0; s=0; np=0; per=0; ty=0; sh=0;
col1=0; col2=0; col3=0; col4=0; col5=0;
delete;

label
f      = 'Factors'
nv     = 'Variables per Factor'
p      = 'Loadings'
v      = 'Total Variables'
r      = 'Interfactor Correlations'
m      = 'Minimum Loading'
u      = 'Unique Factor'
n      = 'Number of Sample Corr Matrices'
s      = 'Sample Size'
np     = 'Number of PA Permutations'
per    = 'Percentile'
ty     = 'Model Type'
sh     = 'Normal or Data Permutation'
col1   = 'Iteration Index'
col2   = 'SEscree'
col3   = 'MAP'
col4   = 'PA'
col5   = 'KG';
run;

data sim.time;
time1 =0;
delete;
run;

%global numcond condnum v;

```

```

%let numcond=0; * Initialize variable to count the overall number of
conditions to be run;
%let condnum=0; * Initialize variable to count the current conditions
being run;
/*****
**                                MAIN SIMULATION    ROUTINE
**
*****
*****/
%do cond=1 %to 2; %*Constant;          %** Used to find number of
conditions apriori (Do Not Change);

    %do f_=1 %to 1;                    %** Number of factors;
    %do nv_=3 %to 15;                  %** Number of variables per
factor;
        %do p1_=5 %to 5;               %** Factor loadings;
        %do r1_=0 %to 4 %by 2;         %** Interfactor correlations;
        %do m1_=3 %to 3;               %** Minimum Saturation (Set
equal to Factor Loading for Constant Saturation);
        %do u_=1 %to 2;                %** No Unique Factor(1)
| Unique Factor(2);
        %do n_=100 %to 100; %*Constant; %** Number of sample
correlation matrices;

        %do np_=50 %to 50; %*Constant; %** Number of Population
Iterations for PA;
        %do per_=95 %to 95;            %*Constant; %** Percentile cutoff for PA;
        %do ty_=1 %to 2;               %** Principal Components(1) |
PAF(2);
        %do sh_=1 %to 1; %*Constant; %** Random normal(1) | Shuffle raw
data(2);

        %if %sysfunc(strip(&nv_)) eq %quote(4) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(6) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(7) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(9) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(11) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(12) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(13) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;
        %if %sysfunc(strip(&nv_)) eq %quote(14) %then %let
nv_=%eval(&nv_+1); %* Skip specific values;

        %let s_=250; %* Specific sample value;

        %let p_=&p1_/10; %* Convert to decimals;
        %let r_=&r1_/10; %* Convert to decimals;
        %let m_=&m1_/10; %* Convert to decimals;

```

```

        %if &cond=2 %then %do; /* Run macro only after counting all
conditions;

        %let seed=0; /* Set the global seed value ***** Important!!
*****;

        %let condnum=%eval(&condnum+1); /* Keep track of current condtion
number;
        %GENCORR(&f_,&nv_,&p_,&r_,&m_,&u_,&n_,&s_,&np_,&per_,&ty_,&sh_);
/* Actual macro call;

        data work.results; /* Merge all condition variables with results
variables;
        set work.results;
        f=&f_; nv=&nv_; p=&p_; v=&v; r=&r_; m=&m_; u=&u_; n=&n_; s=&s_;
np=&np_; per=&per_; ty=&ty_; sh=&sh_;
        run;

        data sim.results; /* Concatenate Results */
        set sim.results work.results;
        run;

        data sim.time; /* Concatenate Time Calculations */
        set sim.time work.time;
        run;

        proc datasets nolist lib=work; /* Clean Up Datasets */
        delete results rout rout1 rout2;
        run;
        %end;
        %else %let numcond=%eval(&numcond+1); /* Count the overall number
of conditions to be run;

        %end;
        %end;
        %end;
        %end;
        %end;
        %end;

        %end;
        %end;
        %end;
        %end;
        %end;

        %end;

        %mend;

        ***** Main Macro Area *****;

        %main; /* Runs the main loop */

```

```

/***** Compute data for time summary *****/
data sim.time;
set sim.time end=last;
if _n_=1 then timeIni=time1;
if last then timeFin=time2;
run;
data work.temp1;
set sim.time;
if timeIni=. then delete;
drop timeFin time1 time2 t1 t2 d1 d2 time_h time_m time_s;
run;
data work.temp2;
set sim.time;
if timeFin=. then delete;
drop timeIni time1 time2 t1 t2 d1 d2 time_h time_m time_s;
run;
data work.temp1;
set work.temp1;
merge work.temp2;

time1=timeIni;
time2=timeFin;

drop timeIni timeFin;

time_s=INTCK( 'second', time1, time2 ) ;
time_m=int(time_s/60) ;
time_h=int(time_m/60) ;

time_s=time_s-(time_m*60);
time_m=time_m-(time_h*60);

d1=datepart(time1);
d2=datepart(time2);
t1=timepart(time1);
t2=timepart(time2);
format d1 d2 worddate16.;
format t1 t2 timeampm13.;

it ='Total';

run;

data sim.time;
set sim.time work.temp1;
drop time1 time2 timeIni timeFin;
label
it          = 'Condition'
d1          = 'Start'
t1          = 'Time'
d2          = 'Finish'
t2          = 'Time'
time_h      = 'Hr(s)'
time_m      = 'Min(s)'
time_s      = 'Sec(s)';

```

```
if it ne 'Total' then it = _n_;  
run;  
  
option notes;  
  
data sim.results (compress=binary); * Compress Results;  
set sim.results;  
run;  
  
dm 'ENDSAS';  
run;  
  
QUIT;
```