University of Rhode Island

DigitalCommons@URI

Open Access Dissertations

1983

A Comparison of Five Rules for Determining the Number of Components in Complex Patterns

William R. Zwick

University of Rhode Island

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Recommended Citation

Zwick, William R., "A Comparison of Five Rules for Determining the Number of Components in Complex Patterns" (1983). *Open Access Dissertations*. Paper 950.

https://digitalcommons.uri.edu/oa_diss/950

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

A COMPARISON OF FIVE RULES
FOR DETERMINING THE NUMBER OF
COMPONENTS IN COMPLEX PATTERNS

BY

WILLIAM R. ZWICK

A DISSERTATION SUBMITTED
IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

1983

#112156.0

ABSTRACT

The performance of five methods for determining the number of components to retain (Horn's parallel analysis, Velicer's MAP, Cattell's SCREE, Bartlett's test and Kaiser's eigenvalue greater than unity) was investigated across seven systematically varied factors (sample size, number of variables, number of components, component saturation, equal or unequal numbers of variables per component, and the presence or absence of unique and complex variables). Five sample correlation matrices were generated at each of two levels of sample size from the 48 known population correlation matrices representing six levels of component pattern complexity. The performance of the parallel analysis and the MAP methods was generally the best across all situations. The SCREE test was generally accurate but variable. Bartlett's test was less accurate and more variable than the SCREE test. Kaiser's method tended to severely overestimate the number of components. Recommendations concerning the conditions under which the methods are accurate and the most effective and useful applications of combinations of methods are discussed.

ACKNOWLEDGMENT

First and foremost I would like to thank my father for his encouragement and quiet confidence. His assumption of my success has made many goals possible.

Dr. Wayne Velicer is a skilled and generous teacher. I wish particularly to thank him for imparting the desire as well as the means to attempt this project. I wish to thank my committee as a whole for their instruction and support through the stages of this research. I wish also to thank Dr. Peter Merenda for his freely given advice, guidance and expert judgment.

Special thanks to Clarice Coleman for her skilled typing and Job-like acceptance of my many weaknesses as an author. I would also like to note the excellent resources made available to me by the staff of the Academic Computer Center of the University of Rhode Island. Their support of my research at the University has been outstanding.

Finally, I would like to thank Fran Danowski for the encouragement and warm friendship she shared with my daughter and me.

PREFACE

Determining the number of principal components to retain has been a troublesome applied research problem. There are many methods to choose from and little compelling evidence of any one rule's superiority. Dr. Velicer's work on a particular method spurred my interest in the question of examining the accuracy of any such method. My own masters thesis, under his guidance, convinced me there was more to learn about the strengths and weaknesses of commonly applied methods.

This dissertation presents, in manuscript form, my continued efforts to examine this question. I have attempted to identify the variables which effect the performance of such methods and have tried to provide evidence of the conditions under which a method may be expected to be most accurate.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENT.	iii
LIST OF TABLES	viii
Chapter	
I. INTRODUCTION.	1
Principal Component Analysis	1
Factor Analysis	2
Selection of Techniques	3
Properties of Retained Components	3
Number of substantial loadings	4
Variance accounted for	4
Component reliability.	4
Determining the Number of Components.	6
Bartlett's test (BART)	6
Eigenvalue greater than 1.0 (K1)	7
The minimum average partial (MAP).	9
The SCREE test	9
Parallel analysis (PA)	11

TABLE OF CONTENTS (CONTINUED)

	<u>Page</u>
Variables Affecting Decision Methods.	12
Sample size.	13
Number of variables.	13
Component saturation	13
Component identification	13
Special variables.	14
Methods To Be Included.	14
 II. METHOD.	 15
 Method of Data Generation	 15
Procedure	16
Data Generation	19
 III. RESULTS	 21
 IV. DISCUSSION.	 28
 V. REFERENCES.	 39

TABLE OF CONTENTS (CONTINUED)

	<u>Page</u>
APPENDICES	
A. DETERMINING THE NUMBER OF COMPONENTS.	55
B. DETAILED ANALYSIS	67
C. CONSTRUCTION OF POPULATION MATRICES	103
BIBLIOGRAPHY	106

LIST OF TABLES

	<u>Page</u>
1. The Relationship Between the Number of Substantial Component Loadings and Eigenvalue Size to Major, Minor and Trivial Components.	44
2. Number of Variables Loading Substantially on Each Component Under the Condition of Unequal Variables/Component for Different Values of P and MJC	45
3. Correlations of the Trained Scree Raters with Each Other at Each Level of the Number of Variables, Pattern Complexity and Component Saturation	46
4. Means and Standard Deviations of the Difference from the Population MJC at Different Sample Sizes when P = 36. . .	47
5. Means and Standard Deviations of the Difference from the Population MJC at Different Sample Sizes when P = 72. . .	48
6. Means and Standard Deviations of the Difference from the Population MJC at Different Numbers of Variables Per Component when P = 36	49
7. Means and Standard Deviations of the Difference from the Population MJC at Different Numbers of Variables Per Component when P = 72	50
8. Means and Standard Deviations of the Difference from the Population MJC at Different Levels of Pattern Complexity when P = 36 and Saturation is .5	51

	<u>Page</u>
9. Means and Standard Deviations of the Difference from the Population MJC at Different Levels of Pattern Complexity when P = 36 and Saturation is .8	52
10. Means and Standard Deviations of the Difference from the Population MJC at Different Levels of Pattern Complexity when P = 72 and Saturation is .5	53
11. Means and Standard Deviations of the Difference from the Population MJC at Different Levels of Pattern Complexity when P = 72 and Saturation is .8	54
12. Percent of Each Method's Estimate Deviating a Set Amount from the Population MJC when P = 36	55
13. Percent of Each Method's Estimate Deviating a Set Amount from the Population MJC when P = 72	56
B-1. Means and Standard Deviations of the Difference From the MJC with or without Equal Number of Variables per Component (EP/m = Y, N) and Unique Variables (UNIQ = Y, N) at Low and High Component Saturation when P = 36, N = 72, with 12 Variables per Component and Complex Variables Absent .	70
B-2. Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables Per Component (EP/m = Y, N) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation when P = 36, N = 72, with 12 Variables Per Component and Complex Variables Present.	73

- B-3. Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables Per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation when P = 36, N = 72, with
6 Variables Per Component and Complex Variables Absent. . 76
- B-4. Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation when P = 36, N = 72, with
6 Variables per Component and Complex Variables Present . 77
- B-5. Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation when P = 36, N = 180, with
12 Variables per Component and Complex Variables Absent . 78
- B-6. Means and Standard Deviations of the Difference from the MJC
with and without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 36, N = 180, with
12 Variables per Component and Complex Variables Present. 79

- B-7. Means and Standard Deviations of the Difference from the MJC
with and without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 36, N = 180, with
6 Variables per Component and Complex Variables Absent. . 80
- B-8. Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 36, N = 180, with
6 Variables per Component and Complex Variables Present. 81
- B-9. Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 144, with
Variables per Component and Complex Variables Absent. . . 82
- B-10. Means and Standard Deviations of the Difference from the MJC
with and without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 144, with
12 Variables per Component and Complex Variables Present. 83
- B-11. Means and Standard Deviations of the Difference from the MJC
with and without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 144, with
8 Variables per Component and Complex Variables Absent. . 84

- B-12.Means and Standard Deviations of the Difference from the MJC
with and without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 144, with
8 Variables per Component and Complex Variables Present . 85
- B-13.Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 380, with
12 Variables per Component and Complex Variables Absent . 86
- B-14.Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 380, with
12 Variables per Component and Complex Variables Present. 87
- B-15.Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 380, with
8 Variables per Component and Complex Variables Absent. . 88
- B-16.Means and Standard Deviations of the Difference from the MJC
with or without Equal Number of Variables per Component
(EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low
and High Component Saturation and P = 72, N = 380, with
8 Variables per Component and Complex Variables Present . 89

	<u>Page</u>
B-17.Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of Sample Size (N = 72, 180) and Saturation when P = 36.	91
B-18.Means and Standard Deviations of the Difference from the Population Value of MJC when Complex Variables are Absent or Present, Saturation is Low or High and when P = 36.	92
B-19.Means and Standard Deviations of the Difference from the Population Value of MJC when there are Equal or Unequal Numbers of Variables per Component (EP/m = N, Y), Saturation is Low or High and when P = 36	93
B-20.Means and Standard Deviations of the Difference from the Population Value of MJC when Unique Variables are Absent or Present (UNIQ = N, Y), Saturation is Low or High and when P = 36	94
B-21.Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of the Number of Variables per Component (#P/MJC = 6, 12) and Saturation when P = 36	95
B-22.Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of Sample Size and Saturation when P = 72.	96

B-23.Means and Standard Deviations of the Difference from the
Population Value of MJC at each Level of the Number of
Variables per Component ($\#P/MJC = 8, 12$) and Saturation
when $P = 72$ 97

B-24.Means and Standard Deviations of the Difference from the
Population Value of MJC when Complex Variables are
Absent or Present, Saturation is Low or High and when
 $P = 72$ 98

B-25.Means and Standard Deviations of the Difference from the
Population Value of MJC when there are Equal or Unequal
Numbers of Variables per Component ($EP/m = N, Y$),
Saturation is Low or High and when $P = 72$ 99

B-26.Means and Standard Deviations of the Difference from the
Population Value of MJC when Unique Variables are Absent
or Present ($UNIQ = N, Y$), Saturation is Low or High and
when $P = 72$ 100

B-27.Percent of Decisions at Different Levels of Deviation from
the Population Value at Low and High Saturation when
 $P = 36$ 101

B-28.Percent of Decisions at Different Levels of Deviation from
the Population Value at Low and High Saturation when
 $P = 72$ 102

Introduction

The representation of a large set of observed variables (P) by a smaller set (m) has been identified as a common problem in the behavioral sciences (Bartlett, 1950, 1951; Guttman, 1954; Horst, 1965; Van de Geer, 1971). Horst (1965) and Van de Geer (1971) discuss principal component analysis (PCA) as an approach to this problem. Another approach is common factor analysis (CFA). Although both PCA and CFA allow a large set of observed variables to be represented by a smaller set, there is disagreement concerning how to determine the number (m) of components or factors required to construct the smaller set. This study presents the results of a Monte Carlo evaluation of five methods for determining m .

Principal Component Analysis

Hotelling (1933) introduced this widely used procedure (Glass and Taylor, 1966; Kaiser, 1970). The first principal component, Y_1 , is defined as the weighted combination of the P observed variables which has the greatest sample variance under the constraint the weight vector is of unit length. Each subsequent principal component Y_j is similarly defined as the weighted combination with maximum variance and unit length weight vector that is orthogonal to all previous components.

The principal component solution may be viewed as an eigen decomposition of the $P \times P$ sample correlation matrix R ,

$$\text{where } R = L'D^2L \quad (1)$$

D^2 is the $P \times P$ diagonal matrix containing the eigen roots of R and L is a $P \times P$ matrix which contains the corresponding eigen vectors.

When component analysis is employed as a parsimony model, only the first m components are retained. For this solution of m principal components, the component pattern (A) may be written as

$$A = L_m D_m \quad (2)$$

where D_m contains the first m eigen roots and L_m contains the corresponding first m eigen vectors. Kaiser (1970) reported on the widespread use of PCA in this manner. Velicer (1974, 1976, 1977) and Velicer, Peacock and Jackson (1982) have shown this use of PCA and CFA result in essentially equivalent solutions.

Factor Analysis

A second class of procedures, called common factor analysis (CFA) has also been employed to express a set of P variables more parsimoniously as a smaller set. The factor analytic model specifies a $P \times P$ correlation (or covariance) matrix may be accounted for by m common and P unique factors. This model may be expressed as

$$R = A A' + U^2 \quad (3)$$

where A is a $P \times m$ pattern matrix and U^2 is the $P \times P$ diagonal matrix of weights for the unique factors. U is conceived of as that part of the item score not "explained" by the common factors. It is important to note that m is frequently assumed to be known for the derivation of these procedures. Sometimes the maximum likelihood test is employed to test if the assumed number of factors is correct.

Selection of Techniques

Since both CFA and PCA are available as data reduction techniques, it is important to note some differences between them. The CFA approach requires that m , the dimension of the reduced set of variables, be known prior to the analysis. The value of m may be determined in one of two general ways. First, some method of determining m may be applied to a PCA solution and the result then used in the factor analysis solution. A second approach uses a maximum likelihood test of the significance to test the fit for different values of m . Unfortunately, many of the methods applied to the PCA solution provide different results from each other and from the maximum likelihood approach. Further, Jackson and Chan (1980) have discussed numerous difficulties within the maximum likelihood approach itself. In addition to these difficulties some doubt has been cast upon the factor analytic model presented in equation (3). An indeterminacy has been identified in the simultaneous estimation of A and U^2 (Guttman, 1954; Schonemann and Wang, 1972; Steigen and Schonemann, 1979). This indeterminacy is inherent in the CFA procedure. In light of difficulties associated with the requirement that m be known a priori, the indeterminacy of the factor model, the widespread use of PCA, and the general comparability of results across the two methods, this study focused on the PCA procedure.

Properties of Retained Components

The comparison of methods to determine the number of components to retain requires a description of the qualities desirable in a retained component. A review of the properties of principal

components, linked with the goal of data summarization, provides such a discription.

Number of substantial loadings. Intuitively, a parsimony application of PCA requires each retained component to contain at least two substantial loadings. Summarizing power is lost unless at least two variables are represented. Algebraic (Anderson and Rubin, 1956) and statistical (Lawley, 1940, Morrison, 1976) examinations of CFA agree that at least three variables are required before the first factor can be identified. Anderson and Rubin (1956) have further demonstrated that each subsequent identifiable factor must contain at least three non-zero loadings. At a sample level, a minimum of at least three significant loadings are required for factor identification. Since complex loadings satisfy this requirement it is not necessary that P be greater than or equal to $3m$.

Variance accounted for. Principal components analysis proceeds from a correlation matrix, a standardized variance-covariance matrix in which the variance of each original variable is equal to 1.0. The variance of each principal component is equal to the eigenvalue of that component. The sum of all P eigenvalues is equal to P , the number of variables. An eigenvalue of 1.0, therefore, accounts for as much variance as that of a single variable. Components with eigenvalues near zero provide no summarizing power. A component with an eigenvalue greater than 1.0 provides more summarizing power than the original variable.

Component reliability. Kaiser (1960) and Kaiser and Caffrey (1965) addressed the issue of component reliability. Noting that a component must be reliable to be useful, Kaiser (1960) reported that the reliability of a component will always be non-negative when the

eigenvalue exceeds 1.0. Horn (1969) noted that this approach to reliability includes all P variables regardless of their component loadings. In applied usage component scores are usually generated as an unweighted sum of those variables with substantial component loadings. Reliability estimates based only on those items contributing to the component score can be quite high even when the component eigenvalue is below 1.0 (Horn, 1969).

The component properties reviewed above can be integrated to present an operational definition of a useful component. Conventional use of PCA as a data reduction technique, combined with algebraic and statistical necessity in CFA dictate that, at the population level, at least three non-zero loadings are required to identify a useful component. At the sample level, three significant and substantial loadings are needed. Invoking again the use of PCA as a data reduction technique coupled with the still useful guarantee of at least non-negative component reliability, retained components are required to have an eigenvalue greater than 1.0. Taking these statements together it is suggested that all components with three or more substantial loadings and an eigenvalue of greater than 1.0 be referred to as major components (MJC). Such MJC components would probably be of interest to most investigators. Components which have less than three substantial loadings but an eigenvalue of 1.0 or greater and components which have more than three substantial loadings but an eigenvalue of less than 1.0 may be of interest to some investigators and will be referred to below as minor (MNC) components. Finally, components with both less than three substantial loadings and an eigenvalue less than 1.0 should never be retained. Such components will be referred to below as trivial (TC)

components. Table 1 represents these operational definitions of major, minor and trivial (MJC, MNC, TC) components.

Given these three categories of components, the performance of various rules for retaining components may be examined.

Determining the Number of Components

A number of rules have been suggested to determine the appropriate number of components to retain (Bartlett, 1950, 1951; Cattell, 1966; Crawford, 1975; Everett, 1983; Horn, 1965; Joreskog, 1962; Kaiser, 1960; Revelle and Rocklin, 1979; Veldman, 1974; Velicer, 1976). These rules often do not give the same results (Anderson, Acito and Lee, 1982; Cattell and Vogelman, 1977; Hakstian, Rogers & Cattell, 1982; Horn, 1965; Linn, 1968; Zwick and Velicer, 1982). Applied researchers are, therefore, often at a loss as to how to proceed. Conflicting research conclusions can be traced to differing methods of defining the correct number of components.

This section will describe the five methods to be evaluated in this study. The methods are: 1) the Bartlett test; 2) the eigenvalue greater than 1.0 rule; 3) the minimum average partial rule; 4) the scree test and 5) the parallel analysis method.

Bartlett's test (BART). Following Lawley's (1940, 1941) test for maximum likelihood factor analysis, Bartlett (1950, 1951) developed a statistical test of the null hypothesis that the remaining $P-m$ eigenvalues are equal. Each component is excluded sequentially from the test until the null hypothesis of equality fails to be rejected. The first m excluded components prior to the retention of the null hypothesis are retained.

Table 1

The Relationship Between the Number of Substantial Component Loadings
and Eigenvalue Size to Major, Minor and Trivial Components.

Variables	Eigenvalue	
	Less Than	Greater Than
Per Component	1.0	1.0
Less		
Than	Trivial	Minor
3	(TC)	(MNC)
Greater		
Than	Minor	Major
3	(MNC)	(MJC)

BART appears sensitive to the number of subjects employed. Gorsuch (1975) argued that as the number increases, the tests of significance become more powerful and, therefore, less and less substantial differences between eigenvalues are found to be significant. This can lead to the retention of more components as a function of the number of subjects, other things being equal. In response to this, Horn and Engstrom (1979) have suggested changing the alpha level at different levels of N . It should be recalled, however, that as the sample size increases the estimates of population eigenvalues will become increasingly accurate. This increased accuracy leads to smaller differences between equal eigenvalues. It should be the case that, within reasonable ranges of sample size, this increased accuracy offsets the increased power of the Bartlett test when the population eigenvalues are actually equal. In such cases, Zwick and Velicer (1982) found the BART test to be somewhat more accurate with relatively larger samples than with small samples.

Eigenvalue greater than 1.0 (K1). Perhaps the most popular, certainly the most commonly employed method, is to retain the components with eigenvalues greater than 1.0. This method is based on one of three lower bounds discussed by Guttman (1954) for the number of components in image analysis. Kaiser (1960) developed this rationale in conjunction with arguments focusing on component reliability and pattern meaningfulness. He stated that "for a principal component to have positive Kuder-Richardson reliability, it is necessary and sufficient that the associated eigenvalue be greater than one" (p. 145) and "the number of eigenvalues greater than one of the observed correlation matrix led to a number of factors

corresponding almost invariably, in a great number of studies, to the number of factors which practicing psychologists were able to interpret" (p. 145). Gorsuch (1974) noted that many users follow Kaiser (1960) and employ the K1 rule to determine the number of components rather than as a lower bound as originally presented. Difficulties associated with this use are noted by Mote (1970) and Humphreys (1964) who argued that rotation of a greater number of components resulted in more meaningful solutions. They imply the relatively blind use of the K1 rule therefore, may sometimes lead to the retention of too few components.

A number of researchers (Browne, 1968; Cattell and Jaspers, 1967; Horn, 1965; Lee and Comrey, 1979; Linn, 1968; Ravelle and Rocklin, 1979; Yeomans and Golder, 1982; Zwick and Velicer, 1982) however, have found the number of components retained by this method often overestimates the known underlying component structure. Gorsuch (1974) and Kaiser (1960) report the number of components retained by K1 is commonly between one third and one fifth or one sixth the number of variables included in the correlation matrix. A Monte Carlo study by Zwick and Velicer (1982) supports this view. This relationship of retained components to the number of variables is detrimental to the accurate estimation of the underlying component structure. The K1 method, although commonly used, is believed by some critics to sometimes underestimate and by many others to sometimes grossly overestimate the number of components, the latter particularly when there are a large (e.g., P greater than 50) number of variables involved.

The Minimum average partial (MAP). Velicer (1976) has suggested a method based on the matrix of partial correlations. The average of the squared partial correlation is calculated after each of the m components has been partialled out. The minimum average of the squared partial correlation indicates the stopping point for this method. That is, when the average squared partial correlation reaches a minimum, the number of components partialled out is the number of components to be retained. Velicer (1976) demonstrated that the average of squared partials will continue to decrease until the residual matrix most closely resembles an identity matrix. After that point, the average squared partial will increase. Using this rule two or more variables would be expected to have high loadings on each retained component. The method is congruent with the factor analytic concept of "common" factors. Velicer (1976) points out the method is exact, can be applied with any covariance matrix and is logically related to the concept of factors as representing more than one variable. In a recent study (Zwick and Velicer, 1982), it was found the MAP rule was more accurate in identifying a known number of components than was the K1 or BART rule.

The SCREE test. Cattell (1966) described a rule based upon a graph of the eigenvalues. The procedure is called the scree test (SCREE). The test appears simple to apply. The eigenvalues are plotted and those falling above a straight line fit through the P - m smaller values are retained. "A basic rationale for the scree test is that the battery of variables is measuring a limited number of factors well and a larger number of trivial, specific and error factors much less well.... The predominant factors account for most of the variance and are large, whereas the other factors are quite

numerous but small.... The substantive factors will be extracted first and the smaller trivial factors will be removed later." (Gorsuch, 1974, p. 152). A number of complications may occur including: (1) gradual slope from lower to higher eigenvalues with no obvious break point in the line; (2) more than one break point in the line; and (3) more than one apparently suitable line can be drawn through the low values. Horn and Engstrom (1979) have noted the underlying similarity of the logic of Bartlett's chi square test and the scree method. Both tests are based on an analysis (one statistical, the other visual) of the essential equality of the remaining eigenvalues.

The scree test has been most effective when strong components are present with little confounding due to error or unique factors. Tucker, Koopman and Linn (1969) found the scree test to be correct in 12 of 18 cases. Cliff (1970) found it to be accurate, particularly if questionable components are included. Cattell and Jaspers (1967) found the test to be correct in 6 of 8 cases, while Cattell and Vogelmann (1977) reported the test to be accurate over 15 systematically differing analyses. Further, Cliff and Hamburger (1967) found more definite breaks with larger ($N = 400$ vs. $N = 100$) sample sizes and Linn (1968) concurred in this conclusion. Zwick and Velicer (1982) found the scree test to be most accurate with larger samples and strong components. They found the scree test to be the most accurate of four methods considered for determining the number of components to retain across many examples of matrices of known, non-complex, structure.

Use of the scree test always involves issues of interrater reliability. Cattell and Vogelmann (1977) and Zwick and Velicer

(1982) have reported good interrater reliability among naive and among expert judges. However, Crawford and Koopman (1979) have reported extremely low interrater reliabilities. The circumstances associated with high and low interrater reliability on the SCREE procedure have not been identified.

Parallel analysis (PA). Parallel Analysis (Horn, 1965), involves a comparison of the obtained, real data eigenvalues with the eigenvalues of a correlation matrix of the same rank and based upon the same number of observations but containing only random uncorrelated variables. This method is an adaptation of the K1 rule. Guttman's (1954) development of upper and lower bounds was based upon population values. Horn (1965) noted that, at the population level, the eigenvalues of a correlation matrix of uncorrelated variables would all be 1.0. When samples are generated based upon such a matrix, however, the initial eigenvalues exceed 1.0 while the final eigenvalues are below 1.0. Horn (1965) suggested that the eigenvalues of a correlation matrix of P random uncorrelated variables, be contrasted with those of the data set in question, based on the same sample size. Components of the matrix of interest which have eigenvalues greater than those of the comparison random matrix would be retained. This approach integrates the reliability and data summarizing emphasis of the population based K1 rule without ignoring the effect of sample size.

Horn (1965) presented one example of PA in a PCA problem. He recommended that the comparison eigenvalues be based upon a number of generated random matrices to avoid major sampling errors in the estimation of the eigenvalues. Although there has been no published systematic examination of the PA method with PCA, Richman (personal

communication, Oct., 14, 1983) reported a series of simulation studies with the method. PA was found to be very accurate when applied to correlation matrices conforming to the formal factor analytic model. He further reported that PA led to retention of too many components when applied to correlation matrices conforming to the middle model described by Tucker, Koopman, and Linn (1969). The method was more accurate in both cases at larger ($N = 500$) than at smaller ($N = 100$) sample sizes.

Humphreys and Montinelli (1975) applied PA to principal axis factor analysis and found the method accurate over a range of examples. Montinelli and Humphreys (1976) developed a regression equation which accurately predicts the eigenvalues of random correlation matrices with squared multiple correlations inserted as the diagonal. Green (1983) utilized this prediction equation to evaluate the performance of factor analysis of binary items. No such prediction equation has been reported for standard correlation matrices.

Variables Affecting Decision Methods

Previously reported research findings on the performance of the decision methods described above indicate that each may be affected by a set of different factors. These factors include sample size, the number of variables, component saturation, component identification and the presence of special types of variables. This study incorporated each of these influences into a simulation study. Some of the considerations related to each of these factors are presented in this section.

Sample size. Depending upon the decision method employed, it is possible the number of subjects may affect the accuracy of the decision about the number of components to retain. Sample size is typically determined both by practical applied considerations and by the need for accurate estimation. The sample size must be large enough to allow an adequate estimation of the relationships between the variables. On the other hand, in applied settings, large samples may be too expensive to be practical.

Number of variables. With the development of computer technology and software, larger and larger correlation matrices have been submitted to PCA. PCAs of personality inventories at the item level, for instance, often involve 80 to 100 variables. Analysis of 200 variable sets is becoming common. Larger numbers of variables have been reported to increase the accuracy of some rules while decreasing it for others (Zwick and Velicer, 1982).

Component saturation. Linn (1968) and Zwick and Velicer (1982) have demonstrated that the underlying component saturation affects decision methods. Underlying components made up exclusively of high loading (e.g., .80) resulted in more retained components, by various decision methods, than did components exclusively made up of lower loadings (e.g., .40).

Component identification. The accurate identification of a component may depend upon the number of variables which have high loadings on that component. Components defined by less than three variables are not capable of identification. The impact of unequal numbers of variables per component is unclear for any of the rules under discussion.

Special variables. Complex variables have a nonzero loading on more than one component. Unique variables have only one non-zero loading and no other variable loads substantially on the same component. Component patterns containing both complex and unequal variables are believed to occur frequently in applied situations (Tucker, et al, 1968). The effect of these types of variables upon the various decision rules is unclear. Complex or unique variables should lead to the retention of more components by K1 and BART and would make SCREE decisions more difficult. The effect of such variables upon MAP and PA has not yet been examined empirically.

Methods To Be Included

The correct determination of the number of components has been identified as a crucial step in the data reduction application of PCA. There continues to be general disagreement concerning the best method to accomplish this step. This study compares the performance of five decision methods on simulated data sets incorporating variables expected to influence each method. The K1 method was included because it is so widely used. The MAP method was included because of its unambiguous solution, its relation to "common factor" concepts and its good performance in a recent study. Bartlett's statistical method (BART) was included because it is the only statistical method appropriate for PCA solutions. The scree test (SCREE) was included because of its apparent simplicity and its reported validity. The parallel analysis method (PA) was included because of its unambiguous solution and its reported accuracy in PCA and CFA. Each of the chosen methods may be differentially affected by several different variables including sample size, the number of

variables, the degree of component identification and saturation, and the presence of unique or complex variables. The robustness of the five rules in question across these variables is the central focus of this study.

Method

Method of Data Generation

Studies of the effectiveness of the various decision methods may be categorized into one of two types. Historically, the more common type of study employed real data representing either new work or "classic" data sets. These studies employed some logical criteria concerning the appropriate number of components and compared the performance of the proposed decision method to the logically arrived at value (e.g., Cattell, 1966; Horn, 1965; Humphreys and Montanelli, 1975; Velicer, 1976). Such studies, in employing an arbitrary logical criterion, may have inaccurately estimated the performance of the decision method in question. More recently, studies of decision rule effectiveness have employed correlation matrices generated from component structures entirely under the control of the investigator (e.g., Anderson, Acito and Lee, 1982; Cattell and Vogelmann, 1977; Tucker, Koopman and Linn, 1969; Zwick and Velicer, 1982). These studies have the advantage of a known criterion against which to measure the performance of the decision methods. They are, however, open to the criticism that the generated matrices, although conforming to a mathematical model, may not well represent real data and thus lead to inappropriate conclusions (Tucker, Koopman and Linn, 1969).

The question of a rule's accuracy cannot be examined without a known criterion. Although logical arguments can be mounted to defend the number of components present in some data sets, these arguments are always open to question. For the assessment of the impact of various conditions upon a rule's accuracy, generated data of a known number of components is preferable. The issue of generalization to real data sets is an important but separate issue which may be independently addressed in the particular way the data is generated. This study employed an approach similar to the "middle model" of Tucher, Koopman and Linn (1969).

Procedure

The number of variables (P) to be employed was set at 36 and 72. These values represent small and moderately large data sets and accommodate constraints imposed by the selection of the number of components to be included. Larger sets of variables have been shown to have a positive impact on MAP, BART (Zwick and Velicer, 1982) and SCREE (Cattell and Vogelmann, 1982) and a negative impact on K1 (Zwick and Velicer, 1982).

The sample sizes (N) chosen were selected to reflect common, applied usage. They were set as a function of the number of variables. The lower N was set at twice the number of variables. The higher N was set at five times the number of variables. The resulting N 's were 72 and 180 in the cases including 36 variables. When 72 variable cases were examined, N 's of 144 and 360 were selected. These appear to include a representative range of sample sizes as reported in applied educational and psychological research. Larger sample sizes have been shown to moderately improve the

performance of the MAP, SCREE and K1 methods (Cattell and Vogelman, 1977; Zwick and Velicer, 1982) and to sometimes improve and sometimes weaken the accuracy of the BART method (Gorsuch, 1975; Zwick and Velicer, 1982).

As described above, major components (MJC) are defined as those with three or more substantial loading and an eigenvalue greater than 1.0. Two types of minor components (MNC) are defined. First, those with three or more substantial loadings and an eigenvalue of less than 1.0. Second, those with less than three substantial loading and an eigenvalue greater or equal to 1.0. Trivial components (TC) are defined as those with less than three substantial loadings and an eigenvalue of less than 1.0.

The number of major components built into the population correlation matrix was 3 and 6 when P was equal to 36 and 6 and 9 when P was equal to 72. These values were chosen to reflect a reasonable range of reported applied usage. They permit a span of an average number of variables per MJC (P/MJC) from 6 to 12. Although this P/MJC is somewhat high, such values are required to permit variability in the number of variables, while still meeting the constraint of at least 3 substantial loadings in each MJC.

The distribution of P/MJC was constrained to be either an equal number of variables per MJC or an unequal number. For the cases where an unequal number existed, the number was symmetrically distributed about the mean number of variables per MJC appropriate for that matrix. That is, if P were 36, and m were 6, the average number of variables per MJC would be 6. When P/MJC was planned to be unequal, the number of variables per component was 8, 7, 6, 6, 5, and

4. Similar distributions for other combinations of P and m are presented in Table 2.

Component saturation (SAT), the magnitude of the loading of the variables on a MJC, was split between a high of .8 and a low of .5. These values bridge much of the applied range and have been shown (Zwick and Velicer, 1982) to differentially effect four of the decision rules under consideration. Within any one matrix, the component loadings on all major components were equal and either high (.8) or low (.5).

For the purposes of this study, unique variables (UNIQ) were defined as variables which do not load at all on either MJC's or TC's in the population structure. Instead, UNIQ variables represent the only variable loading on one type of MNC. Such an MNC has a population eigenvalue of 1.0. Complex variables are defined here as those variables which load substantially on a MJC but also load minimally (.2) on a second type of MNC in the population structure.

Specific combinations of these variables were constructed. Previous work has indicated that N, P and SAT have an impact upon some of the decision rules. At each level of P and SAT component patterns were constructed to evaluate the impact of N, P/MJC and a number of combinations of factors. A "Best Case" set was defined for comparison purposes. This first level of complexity (1) had an equal number of variables per major component, no MNC's or TC's present and thus no complex or UNIQ variables. Five other levels of structural complexity were created for comparison to the "Best Case". This was done by (2) including complex variables to create these MNC's, (3) letting the number of variables per MJC become unequal, (4) including

Table 2

Number of Variables Loading Substantially on Each Component under the

Condition of Unequal Variables/Component for Different Values of P

and Numbers of Major Components (MJC).

P	MJC	P/MJC	Component								
			1	2	3	4	5	6	7	8	9
36	3	12	14	12	10	-	-	-	-	-	-
36	6	6	8	7	6	6	5	4	-	-	-
72	6	12	15	14	13	11	10	9	-	-	-
72	9	8	12	11	10	9	8	7	6	5	4

unique (as many as the number of MJC's) and complex variables together to provide MNC's and finally (6) a level was constructed which included unequal number of variables per major component and both complex and unique variables. The fifth and sixth levels of complexity were included to highlight the possible importance of the very common situation of unequal numbers of variables per MJC.

Data Generation

Population correlation matrices were created for each combination of the 6 (Complexity) x 2 (P) x 2 (SAT) x 2 (m) factors outlined above. Each population correlation matrix was determined as follows: One appropriate population component pattern (A) was created in accordance with the level of the number of variables factor (P), the level of the saturation factor (SAT), the level of the number of components factor (m) and the level of the complexity factor under consideration. Post-multiplying by its transpose (AA') resulted in a matrix R^* ($R^* = AA'$). Substitution of ones in the diagonal of R^* introduced error and produced a population correlation matrix R ($R = R^* + D^2$). The introduction of ones in the diagonal of R raised it to full rank, allowing subsequent analysis. Five sample correlation matrices were generated based on each of these population correlation matrices (Montanelli, 1975) at each level of the number of observations (N) factor. (Appendix C presents an example of a component pattern, an R^* and a final population correlation matrix, R .)

Principal component analysis was then performed on each of the resulting 480 ($6 \times 2 \times 2 \times 2 \times 2 \times 5$) sample correlation matrices. At the time this analysis was performed, the number of components to be retained by each of the four calculable rules (K1, MAP, PA and

BART) was determined. Horn and Engstrom (1979) have suggested that the alpha level of the BART procedure be adjusted to compensate for the increased sensitivity of the test with large sample sizes. Three alpha levels were selected for use with the BART in this study to incorporate Horn and Engstrom's (1979) recommendation. The Bartlett tests were therefore performed at alpha levels of .05 (BA), .001 (BB) and .0005 (BC) in all cases. The PA decision was based on the mean eigenvalues of 50 random correlation matrices at each level of P and N.

The SCREE test was performed on computer generated plots of the eigenvalues of each of the 480 matrices. These plots were examined by two raters trained (Cattell and Vogelman, 1977) in the SCREE method. The two raters were college graduates who had majored in psychology. Although they were trained in the SCREE procedure they were uninformed of its purpose. The raters were also naive to the exact purpose of the experiment and had no prior applied experience with the SCREE test. The graphs were presented to the raters in different mixed orders. If either rater asked a question about a particular plot, both listened to an explanation based on the instructions given by Cattell and Vogelman (1977). Whenever possible, examples from the Cattell and Vogelman (1977) directions were used to clarify questions. Independently, an experienced expert judge, uninformed as to the details of the experiment but fully familiar with the use of the SCREE test, rated one sample from each of the 96 cells.

The judgment required by the SCREE method raises the question of rater reliability. Table 3 presents the interrater reliability estimates of the mean of the raters' decisions corrected for the

number of raters. The reliability estimates are presented at each level of complexity, saturation and the number of variables. The reliability estimates ranged from .61 to 1.00 with a median value of approximately .88.

The correlations of the mean of the raters decisions with the expert's judgment ranged from .60 to .90 across the 6 levels of complexity. The median and mean (Fisher Z transformed) correlation of the averaged rater's decision with the expert's judgment were both approximately .80.

Results

Each decision method leads to an estimate of the number of major components (MJC) to retain. The difference between these rule determined estimates of MJC (RMJC) and the known population value of the number of MJC's (PMJC) was the primary dependent variable in this study. This difference was computed as $\underline{d} = \text{RMJC} - \text{PMJC}$. The mean difference from the criteria is an estimate of the method's accuracy. Positive $\underline{d}$'s, therefore, indicate overestimations while negative $\underline{d}$'s indicate underestimations. A difference of 0 indicates a correct estimation of MJC. The standard deviation of the difference is an indication of the methods' consistency. Smaller standard deviations indicate more consistent estimates of MJC. The mean and standard deviation of $\underline{d}$ for each method, under various conditions, are presented below.

The results are alternately presented first for the $P = 36$ cases and then, in a parallel fashion, for the $P = 72$ cases. Each level of sample size (Tables 4 and 5), number of variables per major component

Table 3

Interrater reliability of the Trained Scree Raters with Each Other at
Each Level of the Number of Variables, Pattern Complexity and Component
Saturation, Correct for the Number of Raters.

Complexity	Saturation	
	.50	.80
P = 36 Variables		
1	.67	.96
2	.96	.95
3	.82	.98
4	.76	.97
5	.65	.91
6	.77	.77
P = 72 Variables		
1	.97	.99
2	.80	.97
3	.95	1.00
4	.93	.81
5	.78	.82
6	.75	.61

(Tables 6 and 7) and pattern complexity (Tables 8, 9, 10 and 11) will be summarized within each level of P and SAT. Tables 12 and 13 present the proportion of each method's estimates of MJC which deviated a set amount from the population value. This representation of the distribution of the estimates is also presented at each level of P and SAT.

Table 4 presents the means (d) and standard deviations (sd) of the difference between each method's estimate of MJC and the known MJC for different sample sizes when $P = 36$ and the component saturation is .5 and .8. Table 4, therefore, summarizes results collapsed across all six levels of pattern complexity and both levels of the number of variables per MJC in order to allow an examination of the individual impact of sample size. Each of the four rows of differences in Table 4 represent 60 observations. Tables 4 through 11 follow essentially the same format. A detailed description will, therefore, be given only for Table 4.

The first row of Table 4 presents the mean difference of each method's estimate of MJC from the known value when P was 36, the saturation was .5 and the sample size was 72. Under these conditions the MAP method produced a mean difference score (d) of -1.08, an underestimation. The PA method produced a much smaller underestimation of -0.05. The SCREE (0.50) and K1 (8.32) methods both overestimated the criterion with K1 providing a very large overestimation. The results given by the Bartlett method were calculated for alpha levels of .05, .001 and .0005, as indicated above. At each level, the Bartlett method led to underestimations (-2.87, -3.92, -3.98 respectively). The Bartlett method retained fewer components as the sensitivity of the test increased at the more

sensitive alpha levels. The standard deviations (sd's) associated with the mean difference scores for each method are presented in row 2. They ranged from 0.70 for the PA method to 1.68 for the MAP method. The third and fourth rows of Table 4 present the mean difference of each method's estimate of MJC from the known value and the standard deviations when P was 36, the saturation was .5 and the sample size was 180. The increase in N from 72 to 180 appeared to have had minor effects on the MAP, PA and SCREE methods. The K1 and BART methods show some improvement at the higher level of N. Rows 5 and 6 of Table 4 present the mean differences and standard deviations for each method when P was 36, the saturation was .8 and the sample size was 72. All the methods showed improved average estimates of the criterion at this higher level of saturation. It should be noted, however, that the standard deviation of the differences increased for all levels of the BART method and, to a lesser extent, for the K1 rule as well. Rows 7 and 8 of Table 4 present the mean differences and standard deviations for each method when the sample size was increased to 180, P was 36 and the saturation was .8. Compared to the results in rows 5 and 6, the larger sample size resulted in more accurate ($d = 0.0$) and consistent ($sd = 0.0$) estimations by MAP and PA methods. The performance of the SCREE and K1 method was not greatly effected. The three levels of the BART method retained more components at the higher sample size. This led to a larger overestimation at BA and a switch from under- to overestimation at BB and BC. The standard deviations at all three levels of BART appear to have been larger at $N = 180$ than at $N = 72$. The K1 method performed slightly better at the higher sample size at both levels of component saturation. BART retained more components

Table 4

Means and Standard Deviations of the Difference from the Population MJC
at Different Sample Sizes when $P = 36$.

		Method						
Sample Size		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
72	d	-1.08	-0.05	0.50	8.32	-2.87	-3.92	-3.98
	(sd)	(1.68)	(0.70)	(1.04)	(1.21)	(1.10)	(0.81)	(0.79)
180	d	-1.17	0.13	0.68	7.30	-1.78	-2.20	-2.27
	(sd)	(1.84)	(0.39)	(0.95)	(1.09)	(1.26)	(0.97)	(0.92)
Saturation = .8								
72	d	0.10	-0.02	0.27	1.77	0.47	-0.48	-0.60
	(sd)	(0.30)	(0.13)	(0.50)	(1.59)	(2.81)	(2.31)	(2.30)
180	d	0.0	0.0	0.23	1.32	1.23	0.68	0.62
	(sd)	(0.0)	(0.0)	(0.52)	(1.44)	(3.51)	(3.27)	(3.23)

at the higher level of sample size at both levels of component saturation. Table 5 parallels Table 4 with P equal to 72. It summarizes the impact of sample size at both levels of component saturation. The MAP and PA methods were again minimally influenced by the sample size change at both levels of component saturation. When the saturation was .5, the SCREE method showed less overestimation at the higher than at the lower sample size. This effect was not apparent when the saturation was .8.

The role of the number of variables is presented from a different perspective in Tables 6 and 7. Table 6 presents a summary of the results for each average number of variables per MJC (P/MJC) when P equaled 36 and the saturation was .5 and .8. At both levels of saturation MAP, PA and SCREE performed more accurately and consistently when the average number of variables per MJC increased from 6 to 12. K1 showed an increased overestimation when P/MJC increased and the saturation was .5 and a decreased overestimation when P/MJC increased and the saturation was .8. The BART method consistently showed a decrease in the number of components retained as P/MJC increased. When the saturation was .5 this resulted in larger underestimations while at a saturation of .8 BART's estimates changed from overestimations to underestimations as P/MJC increased.

Table 7 parallels Table 6 with P equal to 72. The MAP, PA and SCREE methods showed improved performances at the higher level of P/MJC when the saturation was .5. When the saturation was .8 MAP, PA and SCREE showed essentially no improvement at the higher level of P/MJC . The K1 and BART methods showed some improvement at the higher level of P/MJC at both levels of saturation.

Table 5

Means and Standard Deviations of the Difference from the Population MJC
at Different Sample Sizes when $P = 72$.

Sample Size		Method						
		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
144	d	-0.45	0.02	1.16	17.80	-0.43	-1.60	-1.73
	(sd)	(1.00)	(0.57)	(1.30)	(1.29)	(1.16)	(1.24)	(1.15)
360	d	-0.43	0.07	0.46	15.42	0.40	-0.13	-0.22
	(sd)	(1.06)	(0.25)	(0.79)	(1.27)	(0.74)	(0.43)	(0.45)
Saturation = .8								
144	d	0.02	0.0	0.28	2.97	3.88	2.62	2.50
	(sd)	(0.13)	(0.0)	(0.55)	(2.81)	(3.64)	(2.96)	(2.88)
360	d	0.0	0.0	0.31	2.52	5.03	4.10	3.98
	(sd)	(0.0)	(0.0)	(1.03)	(2.59)	(4.31)	(3.89)	(3.78)

Table 6

Means and Standard Deviations of the Difference from the Population MJC
at Different Numbers of Variables Per Component when P = 36.

		Method						
P/MJC		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
6	d	-2.27	0.05	0.65	7.10	-1.50	-2.67	-2.77
	sd	(1.88)	(0.79)	(1.19)	(0.99)	(1.35)	(1.58)	(1.54)
12	d	0.02	0.03	0.53	8.52	-3.15	-3.45	-3.48
	(sd)	(0.13)	(0.18)	(0.75)	(1.10)	(0.44)	(0.53)	(0.57)
Saturation = .8								
6	d	0.08	0.17	0.33	1.68	2.93	2.17	2.10
	(sd)	(0.28)	(0.13)	(0.59)	(1.70)	(2.94)	(2.54)	(2.54)
12	d	0.02	0.0	0.17	1.40	-1.23	-1.97	-2.08
	(sd)	(0.13)	(0.0)	(0.39)	(1.33)	(1.73)	(1.24)	(1.25)

Table 7

Means and Standard Deviations of the Difference from the Population MJC
at Different Numbers of Variables Per Component when $P = 72$.

P/MJC		Method						
		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
8	d	-0.92	0.07	1.02	15.90	-0.17	-1.37	-1.48
	(sd)	(1.28)	(0.61)	(1.22)	(1.45)	(1.21)	(1.36)	(1.32)
12	d	0.03	0.02	0.60	17.32	0.13	0.37	0.47
	(sd)	(0.18)	(0.13)	(1.00)	(1.74)	(0.85)	(0.66)	(0.65)
Saturation = .8								
8	d	0.02	0.0	0.21	3.02	5.15	4.00	3.87
	(sd)	(0.13)	(0.0)	(0.74)	(3.02)	(4.56)	(4.04)	(3.93)
12	d	0.0	0.0	0.38	2.47	3.77	2.72	2.62
	(sd)	(0.0)	(0.0)	(0.89)	(2.34)	(3.27)	(2.80)	(2.73)

Tables 8 and 9 presents a summary of the methods' performance when P was equal to 36 and the saturation was .5 or .8 at each of six levels of pattern complexity (Complexity). The results are collapsed across both levels of N and P/MJC so that each level of Complexity represents 20 observations. As defined above, at Complexity level 1 the MJC's contain equal numbers of variables. There are no unique or complex variables at Complexity level 1 and hence no MNC's or TC's. Level 2 is the same as level 1 except it includes MNC's made up of low complex loadings. Level 3 is the same as level 1 except the number of variables per MJC is unequal across MJC's. Level 4 is the same as level 1 except it includes MNC's made up of unique variables. Level 5 combines both TC's made up of complex loadings and MNC's made up of unique variables. Level 6 is the same as level 5 except the number of variables loading on each MJC is unequal across major components.

The range of pattern complexity affected the methods differently. Although the methods tended to perform best at Complexity level 1, they had different worst cases. When the saturation was .5, in Table 8, the worst cases were: MAP and PA at level 5; SCREE at level 2; K1 and BART at level 4. A comparison of Tables 8 and 9 indicates, MAP, PA, SCREE and K1 showed substantial improvement at all levels of Complexity when the saturation was .8. At this saturation level PA underestimated slightly only at Complexity level 6. MAP overestimated slightly at Complexity levels 2, 5 and 6. SCREE overestimated slightly at all levels of Complexity and level 6 resulted in its largest overestimation. The BART method overestimated slightly or not at all at levels 1, 2 and 3. K1

Table 8

Means and Standard Deviations of the Difference from the Population MJC
at Different Levels of Pattern Complexity when P = 36 and Saturation =
.5.

		Method						
Complexity		MAP	PA	SCREE	K1	BA	BB	BC
1	d	-0.30	0.0	0.38	7.15	-0.90	-1.15	-1.20
	(sd)	(0.66)	(0.32)	(0.92)	(1.46)	(1.29)	(1.46)	(1.51)
2	d	-0.50	-0.10	0.88	7.35	-1.00	-1.40	-1.45
	(sd)	(0.89)	(0.45)	(0.93)	(1.22)	(1.34)	(1.54)	(1.54)
3	d	-0.80	-0.10	0.50	7.30	-1.20	-1.40	-1.55
	(sd)	(1.06)	(0.45)	(1.03)	(1.03)	(1.24)	(1.27)	(1.39)
4	d	-1.60	0.0	0.38	8.45	-1.55	-1.85	-1.90
	(sd)	(2.09)	(0.65)	(0.93)	(1.10)	(1.79)	(1.84)	(1.80)
5	d	-2.05	-0.25	0.58	8.20	-1.30	-1.85	-1.90
	(sd)	(2.50)	(0.79)	(0.96)	(0.95)	(1.59)	(1.93)	(1.89)
6	d	-1.50	0.20	0.85	8.40	-1.30	-1.70	-1.75
	sd	1.96	0.62	1.18	1.10	1.42	1.66	1.68

Table 9

Means and Standard Deviations of the Difference from the Population MJC
 at Different Levels of Pattern Complexity when $P = 36$ and Saturation =
.8.

		Method						
Complexity		MAP	PA	SCREE	K1	BA	BB	BC
1	d	0.0	0.0	0.12	0.15	0.05	0.0	0.0
	(sd)	(0.0)	(0.0)	(0.32)	(0.37)	(0.22)	(0.0)	(0.0)
2	d	0.20	0.0	0.25	0.15	0.30	0.10	0.0
	(sd)	(0.41)	(0.0)	(0.52)	(0.37)	(0.57)	(0.31)	(0.0)
3	d	0.0	0.0	0.10	0.10	0.05	0.0	0.0
	(sd)	(0.0)	(0.0)	(0.26)	(0.31)	(0.22)	(0.0)	(0.0)
4	d	0.0	0.0	0.32	2.90	3.20	2.70	2.50
	(sd)	(0.0)	(0.0)	(0.69)	(0.85)	(1.61)	(1.72)	(1.88)
5	d	0.05	0.0	0.32	2.85	4.55	3.80	3.65
	(sd)	(0.22)	(0.0)	(0.57)	(0.74)	(2.06)	(2.02)	(1.81)
6	d	0.05	-0.05	0.38	3.10	3.45	3.00	2.90
	(sd)	(0.22)	(0.22)	(0.54)	(0.64)	(1.85)	(1.92)	(1.97)

overestimated markedly at Complexity levels 4, 5 and 6 as did BART. Levels 4, 5 and 6 all contain unique variables.

Tables 10 and 11 parallel Tables 8 and 9 with P equal to 72. As was the case when P was 36, the range of complexity appears to have differentially effected the method's performance. At a saturation of .5, in Table 10, MAP was quite accurate at levels 1, 2, 4 and 5 but underestimated erratically at levels 3 and 6. At a saturation of .8, in Table 11, MAP was very accurate at all levels of complexity. PA was quite accurate across all levels of complexity with marked improvement at the .8 saturation level. At that level PA was always accurate. The SCREE method generally overestimated somewhat at each level of complexity. It performed worst when the saturation was .5 at levels 2 and 4 and when the saturation was .8, level 4. The K1 method gave gross overestimates at all levels of complexity when the saturation was .5. It was quite accurate when the saturation was .8 at levels 1, 2 and 3. At the same saturation at levels 4, 5 and 6 the method consistently overestimated the criterion. The BART method showed a moderate range of underestimation when the saturation was .5 with the worst case appearing to be level 6. When the saturation was .8 BART performed well at levels 1 and 3, overestimated moderately at level 2 and overestimated greatly at levels 4, 5 and 6.

A general overview of the performance of the different methods may be gained by calculating the percent of times each method's estimate deviated a set amount from the criterion. Since P and saturation appear to have had the most substantial impact on the methods, the percentages were computed at each level of these variables. Deviations of greater than three were collapsed for

Table 10

Means and Standard Deviations of the Difference from the Population MJC
at Different Levels of Pattern Complexity when $P = 72$ and Saturation =
.5.

		Method						
Complexity		MAP	PA	SCREE	K1	BA	BB	BC
1	d	0.0	0.0	0.82	15.95	-0.25	-0.60	-0.75
	(sd)	(0.46)	(0.0)	(1.09)	(1.82)	(0.55)	(0.82)	(0.79)
2	d	-0.05	0.05	1.18	16.00	-0.15	-0.60	-0.65
	(sd)	(0.22)	(0.22)	(1.08)	(2.10)	(0.37)	(0.68)	(0.67)
3	d	-0.95	-0.20	0.40	15.90	-0.60	-1.00	-1.10
	(sd)	(1.19)	(0.41)	(0.75)	(1.71)	(1.31)	(1.38)	(1.25)
4	d	-0.10	0.30	1.00	17.15	-0.45	-0.95	-1.10
	(sd)	(0.45)	(0.57)	(1.48)	(1.46)	(0.94)	(1.19)	(1.25)
5	d	-0.10	0.20	0.90	17.15	-0.40	-0.85	-1.00
	(sd)	(0.45)	(0.41)	(1.11)	(1.39)	(9.99)	(1.22)	(1.21)
6	d	-1.45	-0.10	0.35	17.50	-0.70	-1.20	-1.25
	(sd)	(1.60)	(0.55)	(0.97)	(1.24)	(1.45)	(1.58)	(1.55)

Table 11

Means and Standard Deviations of the Difference from the Population MJC
 at Different Levels of Pattern Complexity when $P = 72$ and Saturation =
.8.

		Method						
Complexity		MAP	PA	SCREE	K1	BA	BB	BC
1	d	0.0	0.0	0.12	0.0	0.05	0.0	0.0
	(sd)	(0.0)	(0.0)	(0.32)	(0.0)	(0.22)	(0.0)	(0.0)
2	d	0.05	0.0	0.18	0.35	1.70	0.90	0.80
	(sd)	(0.22)	(0.0)	(0.41)	(0.59)	(1.22)	(0.91)	(0.89)
3	d	0.0	0.0	0.0	0.05	0.05	0.0	0.0
	(sd)	(0.0)	(0.0)	(0.0)	(0.22)	(0.22)	(0.0)	(0.0)
4	d	0.0	0.0	1.05	5.40	5.85	5.30	5.20
	(sd)	(0.0)	(0.0)	(1.31)	(0.94)	(2.03)	(2.05)	(1.96)
5	d	0.0	0.0	0.40	5.40	7.95	7.10	6.80
	(sd)	(0.0)	(0.0)	(0.50)	(0.82)	(2.16)	(2.07)	(1.99)
6	d	0.0	0.0	0.0	5.25	7.60	6.85	6.65
	(sd)	(0.0)	(0.0)	(1.08)	(0.85)	(2.76)	(2.56)	(2.64)

simplicity of presentation. Differences of 0 indicate accurate estimates. These percentages are presented in Tables 12 and 13.

As Table 12 indicates, at both levels of saturation when P was 36, PA was clearly the most frequently accurate method followed by MAP and SCREE. K1's tendency to overestimate was marked. The K1 method never underestimated. The BART method was quite inaccurate and variable at both levels of saturation.

Table 13 parallels Table 12 with P equal to 72. Again, PA was the most frequently accurate method at both levels of saturation. Both the PA and the MAP methods showed improved performance when P was equal to 72 compared to 36. The PA and MAP methods were nearly equivalent when the saturation was .8. The K1 method showed essentially the same pattern of results when P was 72 as at 36 because of attenuated range on these tables. The BART method retained more components when P was 72 than 36. BART was more often accurate when the saturation was .5 than .8 when P was 72.

Discussion

The question of interest in this study was the ability of five decision methods to estimate the number of major components present in the population correlation matrices given only the generated sample matrices. The difference between the estimated number and the defined number of major components served as the primary dependent variable in this simulation study. The standard deviation of the difference scores gave further information about each method's consistency. Finally, the percent of correct decisions and the

Table 12

Percent of Each Method's Estimate Deviating a Set Amount from the
Population MJC when $P = 36$.

Deviation	Method						
	MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5							
<u>+</u> 3	0.0	0.0	4.1	100.0	0.0	0.0	0.0
+2	0.0	1.7	16.7	0.0	0.0	0.0	0.0
+1	0.8	10.0	33.3	0.0	1.7	0.0	0.0
<u>±</u> 0	59.2	82.5	36.7	0.0	9.2	2.5	1.7
-1	11.7	2.5	6.7	0.0	19.2	11.7	10.8
-2	9.2	3.3	2.5	0.0	10.0	12.5	12.5
<u>-</u> 3	19.2	0.0	0.0	0.0	59.9	73.3	75.0
Saturation = .8							
<u>+</u> 3	0.0	0.0	0.8	36.7	25.8	17.5	17.5
+2	0.0	0.0	5.0	12.5	1.7	7.5	7.5
+1	5.0	0.0	20.8	7.5	14.2	0.8	0.8
<u>±</u> 0	95.0	99.2	73.3	43.3	26.7	34.2	33.3
-1	0.0	0.8	0.0	0.0	6.7	6.7	6.7
-2	0.0	0.0	0.0	0.0	6.7	7.5	4.2
<u>-</u> 3	0.0	0.0	0.0	0.0	18.3	25.8	30.0

Table 13

Percent of Each Method's Estimate Deviating a Set Amount from the
Population MJC when P = 72.

Deviation	Method						
	MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5							
+3	0.0	0.0	9.2	100.0	1.7	0.0	0.0
+2	0.0	0.8	19.1	0.0	5.8	0.8	0.0
+1	2.5	8.3	21.7	0.0	10.8	1.7	1.7
+0	75.8	85.8	46.7	0.0	58.3	47.5	41.7
-1	6.7	4.2	2.5	0.0	16.7	25.8	31.7
-2	6.7	0.0	0.8	0.0	5.8	11.4	10.8
-3	8.3	0.0	0.0	0.0	0.8	12.5	14.1
Saturation = .8							
+3	0.0	0.0	5.0	50.0	60.0	47.5	45.8
+2	0.0	0.0	1.7	0.8	3.3	8.3	9.2
+1	0.8	0.0	22.5	5.0	9.2	3.3	3.3
+0	99.2	100.0	69.2	44.2	27.5	40.8	41.7
-1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-2	0.0	0.0	0.8	0.0	0.0	0.0	0.0
-3	0.0	0.0	0.8	0.0	0.0	0.0	0.0

percent of decisions at specified levels of deviation from the criterion were also calculated.

The five decision rules employed were the eigenvalue greater than one rule (K1), Bartlett's test (BART), the scree test (SCREE), the minimum average partial method (MAP) and the parallel analysis method (PA). The performance of the five methods for determining the number of components was examined in ten samples drawn from each of 48 simulated population correlation matrices over a range of component pattern complexity. The least complex pattern replicated earlier work (Zwick and Velicer, 1982) and included only equal numbers of variables per component and no unique or complex variables. Component pattern complexity was varied by modifying this clear, least complex case with combinations of complex variables, equal and unequal numbers of variables per component, and unique variables. The resultant six levels of complexity are felt to cover a wide enough range to permit a test of the relative strengths and weaknesses of the decision methods examined. Major components (MJC) were defined as those having more than three substantial loadings and an eigenvalue greater than or equal to 1.0 at the population level. Two types of minor components (MNC) were defined. It is felt this patterns expand upon the formal model and incorporates cases likely to be encountered in real data analyses.

The K1 rule was found to overestimate consistently the number of major components. It never underestimated. This finding is consistent with those of Cattell and Jasper (1967), Linn (1968), Yeomans and Golden (1982) and Zwick and Velicer (1982). At a component saturation of .5, the number retained often fell in the 1/3 to 1/2 of P range discussed by Gorsuch (1974). As the number of

variables increased, so did the number of components retained. K1 retained more components when unique variables were included in the population pattern. These findings are clearly contrary to those of Humphreys (1964) and Mote (1970). In their examination of actual data sets they concluded the K1 method sometimes retained too few components. It would appear that either their data represented a type of component complexity not included in this study or their original judgments of the number of components in their data sets were overestimates. Given the apparent functional relationship of the number of components retained by K1 to the number of variables and to the presence of unique variables, and the repeated reports of the method's inaccuracy, the K1 rule cannot be recommended for PCA. It certainly has not been supported as the best automatic choice as is the case in a number of currently available statistical packages.

The BART method's performance was the most variable of those examined. In addition to variability, the method was sensitive to a number of influences. Increases in N , P and SAT as well as the use of conservative alpha levels and the presence of unique variables all lead to the retention of more components. The first four of these influences may be seen as affecting the statistical power of the Bartlett test. In data sets where the $P-m$ eigenvalues were in fact equal at the population level, Zwick and Velicer (1982) found the method to be moderately accurate. In the broader range of complexity examined here, the test tended to retain both types of minor components defined above. Although examination of different alpha levels led to fewer or greater numbers of components retained, the accuracy and consistency of the method did not appear to be markedly improved by adjusting alpha levels to sample size (see Table 4).

Other factors present in this study appear to have had greater influence on the performance of BART, across alpha levels, than did sample size alone.

The Bartlett test is accurate in answering statistical questions concerning the equality of eigenvalues (Bartlett, 1950; 1951). Researchers inclined to examine minor components, particularly early in the course of exploratory analysis, may find the method helpful. However, the Bartlett test cannot be recommended as a general method of determining the number of major components to retain.

The SCREE method had moderate overall reliability when the mean of two trained raters was used. The correlation of the mean of those raters' decisions with an expert judge indicated fair overall agreement. Reports of rater reliability on the SCREE have ranged from very good (Cattell and Jaspers, 1967) to quite poor (Crawford and Koopman, 1979). This range may reflect either the training or the task complexity across research projects. The raters in this study appear to have shown greater agreement at higher than at lower component saturation levels. They also appear to have shown greater agreement when there were more rather than fewer variables. Perhaps more importantly, the interrater reliability of the SCREE procedure had a fairly wide range across levels of complexity and saturation. The moderate reliability of the SCREE method is very problematic for the applied researcher. Unreliability at this point in the analysis may well expose a study to otherwise avoidable experimenter bias. In any case, applied researchers should note that reliability questions always arise in any use of the SCREE method.

In general the SCREE method was more accurate and less variable than either the K1 or BART method. The method was more accurate and

less variable at the higher level of component saturation. Larger sample sizes also improved its accuracy when P was 72 and SAT was .5. Sample size did not appreciably affect SCREE at other levels of P or SAT. This effect of larger sample size is consistent with those reported elsewhere (Cliff and Pennell, 1967; Linn; 1968 and Zwick and Velicer, 1982). The accuracy of the SCREE method was not affected by an increase in the number of variables examined. An increase in the average number of variables per component did not affect the method's accuracy. In an earlier study (Zwick and Velicer, 1982), utilizing non-complex matrices, the SCREE method performed better than MAP when the major components contained 6 or less variables and the saturation was low. This trend can be observed again under more complex conditions.

The SCREE method tended to overestimate rather than to underestimate when it deviated from the criterion value. As reflected in Tables 12 and 13, the SCREE was accurate about 57% of the time. When the SCREE method was in error, 90 percent of the errors were overestimations. The SCREE method appeared to be most variable at the low level of component saturation or when unique or complex variables were present. Nevertheless, even given its variability and tendency toward overestimation, the SCREE method seldom led to the retention of more than two components over the criterion value. Many experienced investigators routinely examine 1, 2 or 3 components above and below the estimate given by their favorite decision method. This practice, coupled with good judgment concerning interpretability may often result in appropriate solutions when the SCREE method is employed. This optimistic assertion notwithstanding, the ever present question of rater reliability, the

tendency to overestimate and the apparent increased variability in the common case of unique or complex variables all argue against the exclusive use of the SCREE method. Given these drawbacks and the availability of at least one clearly superior method, SCREE can no longer be recommended as a stand alone method of choice for determining the number of components in PCA.

In general, the MAP method was more often accurate and less variable than the K1, BART or SCREE methods. It showed an overall tendency to underestimate the criterion. The MAP method was most accurate at the higher level of component saturation or when the average number of variables per component was large. Its performance was not markedly influenced by sample size within the range examined in this study. The MAP method was quite accurate under many conditions and, when accurate, showed little variability. In cases representing both a low level of saturation and a low number of variables per component, the MAP method consistently underestimated the number of major components. This effect can be most clearly seen in Table 6 by comparing the two levels of P/MJC when $P = 36$ and $SAT = .5$. Additional information about this effect can be gleaned from Table 7. The MAP method gave larger underestimates and displayed greater variability when unique variables were present (levels 4, 5 and 6) and when there was an unequal number of variables per component (level 3 and 6). In this simulation study unique variables reduced the number of variables per component by 1 and the presence of unequal numbers of variables per component independently reduced the number of variables per component on the trailing major components. The presence of complex variables (level 2 and 5) probably lowers the major component saturation at the sample level.

Complex variables thus independently increase the effect of low component saturation. The combination of these influences appears to result in components at the sample level which account for less variance than those components containing only a unique variable. MAP apparently halts at the unique variable component. It, therefore, fails to pick up the less well identified major components.

Overall, the MAP method was accurate more often than were the K1, BART or SCREE methods. The MAP method gave results within ± 1 of the criterion between 72% ($P = 36$, $SAT = .5$) and 100% ($P = 36$, $SAT = .8$; $P = 72$, $SAT = .8$) of the time. When it was in error, the MAP method tended to underestimate. Approximately 90% of the MAP errors were underestimations.

The MAP method provides an unequivocal stopping point. It is tied to the concept of parsimony by directly rejecting components identified by only one variable. MAP showed a tendency to underestimate the known number of major components at the component saturation level of .5 when up to six variables loaded on a component. It is quite accurate when component saturation is high or when there is an average of 8 or more variables per component.

Researchers wishing to ignore relatively small major components may wish to use MAP as a primary method of determining the number of components to retain. Examination of one less and up to 2 more components than determined by MAP is recommended.

The PA method was consistently accurate. It was typically the most accurate method at each level of complexity examined. Under no condition examined did the average deviation of PA from the criterion exceed 0.30. The difference scores of the PA method showed less variability than those of any other rule. Increases in sample size,

component saturations and P/MJC improved the PA method's performance when there was room for improvement. It might have been expected that the PA method would overestimate in the presence of minor components made up of unique or complex variables. Some evidence of this is present at $P = 72$, $SAT = .5$, complexity levels 2, 4 and 5. This pattern is not replicated at other levels of SAT or P.

Overall the PA method was the most frequently accurate method examined. It gave results within ± 1 of the criterion for between 95% ($P = 36$, $SAT = .5$) and 100% ($P = 36$, $SAT = .8$; $P = 72$, $SAT = .8$) of the cases examined. When the PA method was in error it showed a slight tendency toward overestimation. Approximately 65% of the PA method's errors were overestimations. The accuracy of the PA method in this study is consistent with the CFA results reported by Humphreys and Montanelli (1975).

Given its overall accuracy and unambiguous solution, the PA method can be recommended as a primary method for determining the number of major components to retain in PCA. Since the PA method showed a slight tendency to overestimate the criterion, a conservative researcher may do well to examine the range of solutions defined by MAP and PA. This study's results indicate that the appropriate number of major components to retain will very likely lie in the set bracketed by the MAP and PA estimates. When MAP and PA do diverge some form of rotational criteria may serve to aid in a final decision.

A major drawback in the applied use of the PA method is the necessity of generating a large set of random correlation matrices at the particular combination of P and N under consideration. The prediction equation developed for principal axis factor analysis

applications of the PA method (Montinelle and Humphreys, 1976) is not appropriate for principal components analysis.

The present study utilized simulated component patterns to compare five methods representing different approaches to the number of components problem. One of these methods, PA, had not been systematically examined in a PCA application before. The five different decision methods were differentially affected by the variables examined. Increases in the number of variables and the presence of unique and complex variables had detrimental effects on K1 and BART. The BART method also appeared to retain more components at larger sample size and higher levels of component saturation. The SCREE method appeared to overestimate and to result in relatively high levels of variability. The presence of unique variables may have increased these effects at high levels of saturation. The MAP method tended to underestimate the number of major components in complex matrices. This effect was clearest when the component saturation was .5 and the number of variables per component was low. The PA method was quite accurate under the conditions examined. It showed a slight overall tendency to overestimate the number of components.

Although previous studies have examined subsets of these rules under some of the conditions examined here, the present study provides comparison across a wider variety of situations than previous investigations. In those areas where the simulated situations were similar, the results of Linn (1968), Humphreys and Montanelli (1975), Cattell and Vogelman (1977) and Zwick and Velicer (1982) were confirmed and expanded.

Further examinations of SCREE, MAP and PA, under more complex conditions are called for. Applied data sets often result in component patterns with average salient saturations in the .3 to .4 range. Such data also often include variables which load substantially on more than one major component. Further developments in the area of rotational criteria may argue for a systematic comparison of those methods with at least SCREE, MAP and PA.

Within the limitations imposed by the simulation approach, the results of this study, paired with previously reported work, permit some conclusions concerning methods of determining the number of components in real data sets. There is no evidence supporting the continued use of K1 or BART as exclusive, primary methods to determine the number of major components to retain. The SCREE procedure has been reported to be relatively accurate. This study is consistent with those reports but indicates that the method is too variable and too likely to overestimate to be employed as a stand alone decision method. The MAP method was generally quite accurate and consistent when the component saturation was high or the component was defined by more than six variables. Outside these conditions it tended to underestimate the criterion. The PA method was consistently the most accurate of the methods examined. Since PA showed a slight tendency to overestimate, it may prove useful to use it in conjunction with MAP. The general application of the PA method is difficult to recommend because programs needed for its application are not widely available.

References

- Anderson, A. D., Acito, F. & Lee, H. (1982). A simulation study of three methods for determining the number of image components. Multivariate Behavioral Research, 17, 493-502.
- Anderson, T. W. & Rubin, H. (1956). Statistical inference in factor analysis. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. The British Journal of Psychology, 3, 77-85.
- Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. The British Journal of Psychology, 4, 1-2.
- Browne, M. W. (1968). A note on lower bounds for the number of common factors. Psychometrika, 33, 233.
- Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.
- Cattell, R. B. & Jaspers, J. (1967). A general plasmode for factor analytic exercises and research. Multivariate Behavioral Research Monographs, 3, 1-212.
- Cattell, R. B. & Vogelman, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. Multivariate Behavioral Research, 12, 289-325.
- Cliff, N. & Hamberger, C. D. (1967). The study of sampling errors in facator analysis by means of artificial experiments. Psychological Bulletin, 68, 430-445.
- Cliff, N. & Pennell, R. (1967). The influence of communality, factor strength and loading size on the sample characteristics of factor loadings. Psychometrika, 32, (3), 309.

- Crawford, C. B. (1975). Determining the number of interpretable factors. Psychological Bulletin, 82, 2, 226-237.
- Crawford, C. B. & Koopman, P. (1979). Note: Inter-rater reliability of scree test and mean square ratio test of number of factors. Perceptual and Motor Skills, 49, 223-226.
- Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. Multivariate Behavioral Research, 18, 197-218.
- Glass, G. V. & Taylor, P. A. (1966). Factor analytic methodology. Review of Educational Research, 36, 566-587.
- Gorsuch, R. L. (1974). Factor Analysis. Philadelphia: Saunders.
- Green, S. B. (1983). Identifiability of spurious factors using linear factor analysis with binary items. Applied Psychological Measurement, 7, (2), 139-147.
- Guttman, L. (1954). Some necessary conditions for common factor analysis. Psychometrika, 19, 149-162.
- Hakstian, A. R., Rogers, W. T. & Cattell, R. B. (1982). The behavior of number of factors rules with simulated data. Multivariate Behavioral Research, 17, 193-219.
- 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, 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, 283-300.
- Horst, P. (1965). Factor Analysis of Data Matrices. New York: Holt, Rinehart and Winston.

- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24, 417-441, 498-520.
- Humphreys, L. G. (1964). Number of cases and number of factors: An example where N is very large. Educational and Psychological Measurement, 24, 457.
- Humphreys, L. G. & Montanelli, R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. Multivariate Behavioral Research, 10, 193-206.
- Jackson, D. N. & Chan, D. W. (1980). Maximum-Likelihood estimation in common factor analysis: A cautionary note. Psychological Bulletin, 88, 502-508.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.
- Kaiser, H. F. (1970). A second-generation little jiffy. Psychometrika, 35, 401-415.
- Kaiser, H. F. & Caffrey, J. (1965). Alpha factor analysis. Psychometrika, 30, 1-44.
- Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. Proceedings of the Royal Society of Edinburgh, 60, 64-82.
- Lee, H. B. & Comrey, A. L. (1979). Distortions in a commonly used factor analytic procedure. Multivariate Behavioral Research, 14, 301-321.
- Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. Psychometrika, 33, 37-71.

- Montanelli, R. G., Jr. (1975). A computer program to generate sample correlation and covariance matrices. Educational and Psychological Measurement, 35, 195-197.
- Montanelli, R. G. & 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.
- Mote, T. A. (1970). An artifact of the rotation of too few factors: Study orientation vs. trait anxiety. Revista Interamericana de Psicologia, 37, 61-91.
- 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.
- Schonemann, P. H. & Steiger, J. H. (1978). On the validity of indeterminant factor scores. Bulletin of the Psychonomic Society, 12, 287-290.
- Schonemann, P. H. & Wang, M. (1972). Some new results on factor indeterminacy. Psychometrika, 37, 61-91.
- Steiger, J. H. & Schonemann, P. H. (1979). A history of factor indeterminacy. In Shye, S. (Ed.) Theory Construction and Data Analysis in the Behavioral Sciences. San Francisco: Jossey-Bass.
- Tucker, L. B., Koopman, R. F. & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. Psychometrika, 34, 421.
- Van de Geer, J. P. (1971). Introduction to Multivariate Analysis for the Social Sciences. San Francisco: Freeman.
- Veldman, D. J. (1979). Simple structure and the number of factors problem. Multivariate Behavioral Research, 14, 191-200.

- Velicer, W. F. (1974). A comparison of the stability of factor analysis, principal component analysis and rescaled image analysis. Educational and Psychological Measurement, 34, 563-572.
- Velicer, W. F. (1976a). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 3, 321-327.
- Velicer, W. F. (1976b). The relation between factor score estimates, image scores and principal component scores. Educational and Psychological Measurement, 36, 149-159.
- Velicer, W. F. (1977). An empirical comparison of the similarity of principal component, image, and factor patterns. Multivariate Behavioral Research, 12, 3-22.
- 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, 371-388.
- 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.
- Zwick, W. R. & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. Multivariate Behavioral Research, 17, 253-269.

APPENDIX A
DETERMINING THE NUMBER OF COMPONENTS

A number of methods have been suggested to determine the appropriate number of components to retain (Bartlett, 1950, 1951; Cattell, 1966; Crawford, 1975; Everett, 1983; Horn, 1965; Joreskog, 1962; Kaiser, 1960; Revelle and Rocklin, 1979; Veldman, 1974; Velicer, 1976). These methods often do not give the same results (Anderson, Acito and Lee, 1982; Cattell and Vogelman, 1977; Hakstian, Rogers & Cattell, 1982; Horn, 1965; Linn, 1968; Zwick and Velicer, 1982). Applied researchers are, therefore, often at a loss as to how to proceed. Conflicting research conclusions can be traced to differing methods of determining the number of components.

Gorsuch (1974) categorized decision methods for determining the appropriate number of components into Statistical, Mathematical and Non-Trivial groups. A fourth category of more recently developed methods may be described as Rotational Criteria (Crawford, 1975; Revelle and Rocklin, 1979; Veldman, 1974). Methods within each of these categories have something to recommend them and will be examined in turn.

Statistical Methods

Bartlett's test. Following Lawley (1940), Bartlett (1950, 1951) suggested a statistical test of the null hypothesis that the remaining eigenvalues are equal. Guttman (1954) has described why such a test might be useful. The sum of the eigenvalues must equal P . The average eigenvalue will, therefore, be 1.0. Either each eigenvalue must be equal to 1.0 or at least one will be greater, and one less, than 1.0. Guttman (1954) states the distribution of the P eigenvalues about the mean is asymmetric such that if a few are somewhat larger than unity, then a greater number will be somewhat

less than unity. Tests of the essential equality of the remaining $P - m$ eigenvalues are based on the fact that a set of $P - m$ eigenvalues can be equal in only one of two ways: 1) either all P eigenvalues are equal to 1.0 or 2) $P - m$ eigenvalues are equal and each is smaller than 1.0. In either case, no component in the set of components with equal eigenvalues can make a data summarization contribution since such a component cannot account for more than the variance accounted for by a single variable.

In Bartlett's test, each component is sequentially excluded from the test until the null hypothesis of equality fails to be rejected. The first m excluded components prior to the retention of the null hypothesis are retained.

Bartlett presents the formula:

$$V = -K \ln B \quad (A-1)$$

$$\text{where: } K = N - (2P + 5)/6 - (2m/3) \quad (A-2)$$

$$\text{and } B = \frac{(m-1)(m+2)(P)}{((m+1) + (m+2))} \quad (A-3)$$

$$\frac{P-m}{P-m}$$

where: N = the sample size

m = the number of components already retained and

P = the number of variables.

The statistic V is distributed as chi square with degrees of freedom equal to: $(P - m - 1) (P - m + 2)/2N$

Bartlett's test appears sensitive to the number of subjects employed. Gorsuch (1975) argued that as the number increases, the tests of significance become more powerful and, therefore, less and less substantial differences between sample eigenvalues are found to be significant. This can lead to the retention of more components as a function of the number of subjects, other things being equal. In response to this, Horn and Engstrom (1979) have suggested changing the alpha level at different levels of N . It should be recalled, however, that as the sample size increases the estimates of population eigenvalues will become increasingly accurate. This increased accuracy leads to smaller differences between equal eigenvalues. It should be the case that, within reasonable ranges of sample size, this increased accuracy offsets the increased power of the Bartlett test when the population eigenvalues are actually equal. In such cases Zwick and Velicer (1982) found the BART test to be somewhat more accurate with relatively larger samples than with small samples.

Mathematical Methods

Eigenvalue greater than 1.0 (K1). Perhaps the most popular, certainly the most commonly employed method, is to retain those components with eigenvalues greater than 1.0. This method represents one of three lower bounds discussed by Guttman (1954) for the number of components in image analysis. Kaiser (1960) further supported this approach by focusing on component reliability and pattern meaningfulness. He stated that "for a principal component to have positive Kuder-Richardson reliability, it is necessary and sufficient that the associated eigenvalue be greater than one" (p. 145) and "the

number of eigenvalues greater than one of the observed correlation matrix led to a number of factors corresponding almost invariably, in a great number of studies, to the number of factors which practicing psychologists were able to interpret." (p. 145). Gorsuch noted that many users follow Kaiser (1960) and employ the K1 rule to determine the number of components rather than as a lower bound as originally presented. Difficulties associated with this use are noted by Mote (1970) and Humphreys (1964) who argued that rotation of a greater number of components resulted in more meaningful solutions. They imply the relatively blind use of the K1 rule therefore, may sometimes lead to the retention of too few components.

A number of researchers (Browne, 1968; Cattell and Jaspers, 1967; Horn, 1965; Lee and Comrey, 1979; Linn, 1968; Ravelle and Rocklin, 1979; Yeomans and Golder, 1982; Zwick and Velicer, 1982) however, have found the number of components retained by this method often overestimates the known underlying component structure. Gorsuch (1974) and Kaiser (1960) report that the number of components retained by K1 is commonly between one third and one fifth or one sixth the number of variables included in the correlation matrix. This relationship of retained components to the number of variables is detrimental to the accurate estimation of the underlying component structure. The K1 method, although commonly used, is believed by some critics to sometimes underestimate and by many others to sometimes grossly overestimate the number of components, the latter particularly when there are a large (e.g., P greater than 50) number of variables involved.

The Minimum Average Partial (MAP). Velicer (1976) has suggested a method based on the matrix of partial correlations. The average of

the squared partial correlation is calculated after each of the m components has been partialled out. The minimum average of the squared partial correlation indicates the stopping point for this method. That is, when the average squared partial correlation reaches a minimum, the number of components partialled out is the number of components to be retained. Velicer (1976) demonstrated that the average of squared partials will continue to decrease until the residual matrix most closely resembles an identity matrix. After that point, the average squared partial will increase. Using this rule two or more variables would be expected to have high loadings on each retained component. In his presentation of the foundation of the minimum average partial method, Velicer (1976) considers the partial correlation equation:

$$r_{ij \cdot y} = \frac{r_{ij} - r_{iy}}{((1 - r_i^2 y) - (1 - r_j^2 y))^{1/2}} \quad (A-4)$$

where i, j represent any two of the P observed variables and y represents a component. With no loss of generality, assume r_{ij} , r_{iy} and r_{jy} to be positive. The value of the partial correlation will decrease as long as the numerator decreases faster than the denominator. The value of the correlation will increase whenever the reverse is true; for example, when r_{iy} is large and r_{jy} is small. Such a situation would occur if a component has a high correlation with only one variable and near zero correlations with the others...". (p. 323) The method is congruent with the factor analytic concept of "common" factors. Velicer (1976) points out the method is exact, can be applied with any covariance matrix and is

logically related to the concept of factors as representing more than one variable. In a recent study (Zwick and Velicer, 1982), it was found the MAP rule was more accurate in identifying a known number of components than was the K1 or BART rule.

Non-trivial Contribution Methods.

The Percent of variance. The percent of total variance accounted for by the extracted components is perhaps the oldest guideline for determining the number of components to retain. Investigators commonly compute the cumulative percent of variance accounted for after each component is extracted. Sometimes components have been retained until the process has accounted for 75, 85, or some other arbitrary percent of the total variance. Typically, extraction is stopped when the arbitrarily large proportion is accounted for and when extraction of the next component would contribute little additional variance. Unfortunately, the arbitrary setting of a goal percentage has neither a firm logical nor a firm mathematical basis. For example, it is useful to note that if P correlated variables are augmented by a set (x) of uncorrelated variables the proportion of variance accounted for by the first component in the P plus x set drops while the meaningfulness of the relationships remains the same. The presence of uncorrelated variables, therefore, directly affects the number of retained components even though the pattern of relationships among the original P variables is unchanged.

Cattell's Scree Test. Cattell (1966) presented a visual rule based upon a graph of the eigenvalues. The procedure, called the scree test, seeks to identify insignificant eigenvalues. The test appears simple to apply. The eigenvalues are plotted and those

falling above a straight line fit through the smaller values are retained. "A basic rationale for the scree test is that the battery of variables is measuring a limited number of factors well and a larger number of trivial, specific and error factors much less well.... The predominant factors account for most of the variance and are large, whereas the other factors are quite numerous but small.... The substantive factors will be extracted first and the smaller trivial factors will be removed later." (Gorsuch, 1974, p. 152). A number of complications may occur including 1) gradual slope from lower to higher eigenvalues with no obvious break point in the line; 2) more than one break point in the line; 3) more than one apparently suitable line drawn through the low values.

Horn and Engstrom (1979) have noted the underlying similarity of Bartlett's chi square test and the scree method. Both tests are based on an analysis (one statistical, the other visual) of the essential equality of the remaining eigenvalues.

The scree test has been most effective when strong components are present with little confounding due to error or unique factors. Tucker, Koopman and Linn (1969) found the scree test to be correct in 12 of 18 cases. Cliff (1970) found it to be accurate, particularly if questionable components are included. Cattell and Jaspers (1967) found the test to be correct in 6 of 8 cases, while Cattell and Vogelmann (1977) reported the test to be accurate over 15 systematically differing analyses. Further, Cliff and Hamburger (1967) found more definite breaks with larger ($N = 400$ vs. $N = 100$) sample sizes and Linn (1968) concurred in this conclusion. Zwick and Velicer (1982) found the scree test to be most accurate with larger samples and strong components. They found the scree test to be the

most accurate of four methods for determining the number of components to retain across many examples of matrices of known, non-complex, structure.

Use of the scree test always involves issues of interrater reliability. Cattell and Vogelman (1977) and Zwick and Velicer (1981) have reported good interrater reliability among naive and among expert judges. Crawford and Koopman (1979) have reported extremely low interrater reliabilities, however.

Parallel analysis. A third non-trivial contribution method based upon an examination of the eigenvalues has been suggested. Parallel Analysis (Horn, 1965), involves a comparison of the obtained, real data eigenvalues with the eigenvalues of a correlation matrix of the same rank and based upon the same number of observations but containing only randomly associated variables. This method is an adaptation of the K1 rule. Guttman's (1954) development of upper and lower bounds was based upon population values. Horn (1965) noted that, at the population level, the eigenvalues of a correlation matrix of randomly associated variables would all be 1.0. When samples are generated based upon such a matrix, however, the initial eigenvalues exceed 1.0 while the final eigenvalues are below 1.0. The smaller the sample size, the more the initial eigenvalues exceed 1.0. Horn (1965) suggested that the eigenvalues of a correlation matrix of P randomly associated variables be contrasted with those of the data set in question, based on the same sample size. Components of the matrix of interest which have eigenvalues greater than those of the comparison random matrix would be retained. This approach integrates the reliability and data summarizing emphasis of the population based K1 rule without ignoring the effect of sample size.

Horn (1965) presented one example of PA in a PCA problem. He recommended that the comparison eigenvalues be based upon a number of generated random matrices to avoid major sampling errors in the estimates of the eigenvalues. Although there has been no published systematic examination of the PA method with PCA, Richman (personal communication, Oct., 14, 1983) reported a series of simulation studies with the method. He found PA to be very accurate when applied to correlation matrices conforming to the formal factor analytic model. He further reported that PA led to retention of too many components when applied to correlation matrices conforming to the middle model described by Tucker, Koopman, and Linn, (1969). The method was more accurate in both cases at larger ($N = 500$) than at smaller ($N = 100$) sample sizes.

Humphreys and Montinelli (1974) applied PA to principal axis factor analysis and found the method accurate over a range of examples. They (Montinelli and Humphreys (1976)) then presented the following regression equation

$$\log Y_i = a_i + b_{Ni} \log (N-1) + b_{Pi} \log \frac{P(P-1) - (i-1)P}{2} \quad (A-5)$$

where i is the ordinal position of the eigen value,

b_{Ni} and b_{Pi} are regression coefficients,

a_i is the intercept,

N is the sample size and

P is the number of variables.

This equation accurately predicts the eigenvalues of random correlation matrices with squared multiple correlations inserted as the diagonal elements. Green (1983) utilized this prediction equation to evaluate the performance of factor analysis of binary

items. No such prediction equation has been reported for standard correlation matrices. Following Montanelli and Humphreys (1976) general rationale but incorporating Bartlett's (1950) presentation concerning degrees of freedom, the following equation is suggested as a starting point to develop a useful prediction equation for PCA.

$$\log Y_i = a_i + b_{N_i} \log (N-1) + b_{p_i} \log \frac{(P-i+2)(P-i-1)}{2} \quad (A-6)$$

Rotational methods

A recent approach to the problem of determining the number of components to retain focuses upon the pattern of loadings which result from the rotation of differing numbers of components. Veldman (1974) and Crawford (1975) emphasized the goal of simple structure in determining the number of components to retain. They have suggested that the rotational criterion used to select the best solution within the set of those available at a given value of m could also serve as a criteria to select the best solution from those available at different levels of m . That is, they suggest one should compare the rotated solution of different numbers of components to find the one which best fits some mathematic definition of simple structure. Veldman (1974) emphasized the orthogonal varimax criteria as his choice while Crawford (1975) presented a more general criteria adaptable to both orthogonal and oblique rotations. Revelle and Rocklin (1979) extended this approach to include the practice of using unweighted component scores and the concept of a minimized residual matrix.

Each of the rotational approaches share the advantage that the retained number of components will provide a best approximation of

simple structure. These approaches emphasize the overall component pattern rather than the properties of any one component. Anderson and Lee (1982) found the varimax criterion to be useful in image analysis approaches but not in PCA. Crawford (1975) reported similar findings.

APPENDIX B
DETAILED ANALYSIS

Sixty-four population correlation matrices were constructed based on component patterns defined by two levels each of six factors (the number of variables ($P = 36, 72$); the component saturation ($SAT = .5, .8$); the number of components (at $P = 36, m = 3, 6$; at $P = 72, m = 6, 9$); the presence or absence of unique ($UNIQ = Y, N$) and complex variables ($CV = Y, N$) and the presence or absence of equal numbers of variables per component ($EP/m = Y, N$)). Five sample correlation matrices were then generated for each population matrix at each of two levels of sample size (at $P = 36, N = 72, 180$; at $P = 72, N = 144, 360$). Principal components analysis was performed on each of the resultant 640 sample correlation matrices. Five methods of determining the number of components (Horn's parallel analysis (PA); Velicer's minimum average partial (MAP); Cattell's SCREE; Bartlett's test (BART) and Kaiser's eigenvalue greater than 1.0 rule (K1)) were applied to each component solution. The performance of the BART was examined at three alpha levels ($BA = .05, BB = .001$ and $BC = .0005$). The difference from the known number of major (MJC) components in the population was calculated for each decision method.

The average difference (d) and the standard deviation of the difference (sd) from the population value for each method on each component pattern are presented in Tables B-1 through B-16. The results for the $P = 36$ and $P = 72$ cases are presented in Tables B-1 through B-8 and B-9 through B-16 respectively. Each row of Tables B-1 through B-16 represents 5 observations.

The individual impact of a number of factors (sample size, presence or absence of complex and unique variables and equal numbers of variables per component and the absolute number of variables per component) on the mean and standard deviation of the difference from

the known criteria are highlighted for each method in Tables B-17 through B-26.

The proportion of each decision method's estimate which deviated set amounts from the known population value is presented in Table B-27 ($P = 36$) and B-28 ($P = 72$). These proportions provide a summary overview of the accuracy and variability of each decision method's performance.

Table B-1 through B-26 follow the same general format. A detailed description is therefore given only for the first table. Table B-1 presents the mean difference and the standard deviation of the difference from the known population value (MTC) for each decision method. The results presented in Table B-1 represented eight generated cases (two levels each of EP/m, UNIQ and SAT fully crossed) where $P = 36$, $N = 72$, there are 12 variables per component ($MJC = 3$) and complex variables are absent. The first row of Table B-1 represents the case in which the component saturation was .5, there were not an equal number of variables per component ($EP/m = N$) and there were no unique variables present ($UNIQ = N$). In that case, the MAP and PA methods showed an average difference (d) of 0.0 from the criterion value. The SCREE (0.70) and K1 (8.60) methods both overestimated, SCREE slightly, K1 markedly. The Bartlett test resulted in moderate ($BA = -0.60$) to somewhat large ($BC = -1.20$) underestimations. The second row of Table B-1 presents the standard deviations of the difference (sd) for that case. The MAP and PA methods showed no variability at all ($sd = 0.0$) in their accurate estimation of the population value. The remaining methods resulted in a range of consistency from 0.0 for Bartlett's test at an alpha level of .001 (BB) to 0.55 for the K1 method and BB. The third and

Table B-1

Means and Standard Deviations of the Difference From the MJC With or Without Equal Number of Variables per Component (EP/m = Y, N) and Unique Variables (UNIQ = Y, N) at Low and High Component Saturation when P = 36, N = 72, with 12 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N	d	0.0	0.0	0.70	8.60	-0.60	-1.00
			(sd)	(0.0)	(0.0)	(0.45)	(0.55)	(0.55)	(0.0)
N	Y	d	0.0	0.0	1.00	9.20	0.0	-0.80	-0.80
			(sd)	(0.0)	(0.0)	(1.17)	(0.45)	(1.22)	(0.84)
	UNIQ	N	d	0.20	0.0	1.00	8.80	-0.20	-.60
			(sd)	(0.45)	(0.0)	(1.27)	(0.84)	(0.45)	(0.55)
Y	Y	d	0.0	0.0	0.90	9.80	-0.60	-1.00	-1.20
			(sd)	(0.0)	(0.0)	(0.55)	(0.84)	(0.55)	(0.71)
Saturation = .8									
	UNIQ	N	d	0.0	0.0	0.30	0.40	0.20	0.0
			(sd)	(0.0)	(0.0)	(0.45)	(0.54)	(0.45)	(0.0)
N	Y	d	0.0	0.0	0.0	2.80	2.40	1.00	1.00
			(sd)	(0.0)	(0.0)	(0.0)	(0.45)	(1.52)	(1.00)
	UNIQ	N	d	0.0	0.0	0.0	0.60	0.60	0.0
			(sd)	(0.0)	(0.0)	(0.0)	(0.55)	(0.89)	(0.0)
Y	Y	d	0.0	0.0	0.10	3.00	2.00	0.80	0.20
			(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(0.71)	(0.84)

fourth rows of Table B-1 present the results for the same case as rows one and two except unique variables were present ($UNIQ = Y$) rather than absent. In this case both the MAP and PA methods continue to be accurate ($d = 0.0$) and consistent ($sd = 0.0$). The other methods showed an increase in the number of components they retained. The SCREE and Bartlett methods also showed an apparent increase in their variability which was not present for the K1 method. Rows five and six present the results when there are an equal number of variables per component ($EP/m = Y$) and unique variables are not present ($UNIQ = N$). The methods range in accuracy from an underestimation of $-.60$ by BB and BC to an overestimation of 8.80 by K1. PA is again accurate ($d = 0.0$) and consistent ($sd = 0.0$). Rows seven and eight parallel the case presented in rows five and six except unique variables are now present ($UNIQ = Y$). The methods continued to range widely in accuracy. The Bartlett test lead to an underestimation of -1.20 (BC) while the K1 method resulted in an overestimation of 9.80 . Both the MAP and PA methods were accurate ($d = 0.0$) and consistent ($sd = 0.0$).

The 9th through 16th rows represent the same cases as rows 1 through 8 except the saturation was increased from $.5$ to $.8$. Under its condition MAP and PA are again very accurate ($d = 0.0$) and consistent ($sd = 0.0$). The SCREE and K1 methods are more accurate than in the saturation = $.5$ cases. Both methods showed reduced overestimations and increased consistency. The Bartlett method retained more components at the higher ($.8$) level of component saturation than at the lower level ($.5$). Under some combinations of conditions this lead to increased accuracy while in other cases it

did not. The variability of the Bartlett method did not appear to be affected by the increased component saturation.

Table B-2 presents the same combination of conditions as Table B-1 except complex variables were present. The methods all performed about as they had when complex variables were absent. The only notable exception appears to have been the retention of a greater number of components by the Bartlett method when the saturation was high and there were an equal number of variables per component and unique variables were present.

The impact of the various conditions upon the performance of each method is presented in detail in Table B-1 through B-16. The most useful understanding of the impact of the conditions may be gleaned by noting the two worst cases for each method. These cases will be presented in order from the most accurate method to the least accurate.

Overall, the PA method was the most accurate. It showed its largest mean deviation ($d = 1.00$) from the criterion when $P = 36$, $N = 180$, $SAT = .5$, there were six variables per component, and complex variables, unique variables and equal variables per component were present (Table B-8). It showed nearly as large an underestimation ($d = -.80$) in the case where $P = 72$, $N = 144$, $SAT = .5$, there were eight variables per component and complex variables, unique variables and equal variables per component were all absent.

The MAP method was the second most accurate method overall. When it was in error it usually underestimated the population value of MJC. The largest mean underestimation ($d = -4.60$) came when $P = 36$, $N = 180$, $SAT = .5$, there were six variables per component and complex and unique variables were present with an equal number of variables,

Table B-2

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables Per Component (EP/m = Y, N) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation when P = 36, N = 72, with 12 Variables Per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N d	0.0	0.0	1.00	9.60	0.0	-1.00	-1.00
		(sd)	(0.0)	(0.0)	(1.27)	(0.55)	(0.71)	(0.0)	(0.0)
N	Y	d	0.0	0.20	0.90	9.60	-.20	-1.00	-1.00
		(sd)	(0.0)	(0.45)	(0.55)	(0.89)	(0.45)	(0.0)	(0.0)
	UNIQ	N d	0.0	0.0	0.70	9.00	0.20	-.80	-0.80
		(sd)	(0.0)	(1.0)	(1.30)	(1.00)	(0.45)	(0.45)	(0.45)
Y	Y	d	0.0	0.20	0.60	9.40	-0.40	-1.00	-1.00
		(sd)	(0.0)	(0.45)	(0.55)	(0.55)	(0.55)	(0.0)	(0.0)
Saturation = .8									
	UNIQ	N d	0.0	0.0	0.40	0.20	1.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.55)	(0.45)	(1.64)	(0.0)	(0.0)
N	Y	d	0.0	0.0	0.10	3.20	2.20	0.80	0.60
		(sd)	(0.0)	(0.0)	(0.22)	(0.45)	(0.84)	(0.45)	(0.55)
	UNIQ	N d	0.20	0.0	0.20	0.20	0.80	0.20	0.0
		(sd)	(0.45)	(0.0)	(0.45)	(0.45)	(1.79)	(0.45)	(0.0)
Y	Y	d	0.0	0.0	0.50	3.20	4.00	2.00	2.00
		(sd)	(0.0)	(0.0)	(0.61)	(0.45)	(1.87)	(0.71)	(0.71)

per component (Table B-8). The second largest mean underestimation by MAP ($d = -3.80$) occurred under the same conditions except complex variables were not present (Table B-7).

The SCREE method was generally fairly accurate but more variable than the PA or MAP methods. When in error, the SCREE method usually overestimated the population value of MJC. The two largest overestimations were $d = 2.70$ and $d = 1.70$ (Tables B-11 and B-12 respectively). Both overestimations occurred when $P = 72$, $N = 144$, $SAT = .5$, there were 8 variables per component, unique variables were present and there were an equal number of variables per component. The largest overestimation occurred with complex variables absent while the second largest occurred when they were present (Table B-12).

The K-1 method generally overestimated the population value of MJC. The largest ($d = 19.4$) and second largest ($d = 19.0$) overestimations are presented in Tables B-9 and B-10 respectively. In both cases $P = 72$, $N = 144$ and $SAT = .5$. The largest overestimations occurred when complex variables were absent and there were not equal numbers of variables per component but unique variables were present. The second largest overestimation occurred when complex variables were present, there was an equal number of variables per component and unique variables were either absent or present ($d = 19.0$ in both cases).

The Bartlett test provided a wide range of estimates, sometimes over and sometimes underestimating the population value of MJC. In any case, the Bartlett method always retained the least components at the most conservative alpha level examined ($\alpha = .005$). The largest overestimation by the Bartlett test ($d = 9.2$) occurred when $P = 72$, $N = 144$, $SAT = .8$, there were 8 variables per component, an

unequal number of variables per component and complex and unique variables were present and the test was conducted at the .05 level (Table B-12). The largest underestimation ($d = -4.8$) occurred when $P = 36$, $N = 72$ with an equal number of 6 variables per component and when complex and unique variables were present. This underestimation occurred at the .001 and .0005 level.

Tables B-17 through B-26 present the individual impact of sample size, the presence or absence of complex and unique variable as well as equal numbers of variables per component and the absolute number of variables per component at each level of P and SAT. Each row of Tables B-17 through B-26 represents 160 observations collapsed across all other factors. In general, component saturation appears to have had the largest impact on each rule. Regardless of other conditions, the methods appeared to perform better at higher than at lower component saturation. A review of worst cases for each method might again be useful.

The PA method showed its largest mean difference ($d = 0.38$) from the population value of MJC when $P = 36$, SAT = .5 and complex variables were present (Table B-18). The same deviation occurred under similar conditions except an equal number of variables were present. The MAP method showed its largest mean deviation ($d = -2.26$) when $P = 36$, SAT = .5 and there were 6 variables per component. The SCREE method's largest deviation ($d = 1.08$) occurred when $P = 72$, SAT = .5 and the sample size was 144. The K1 method's largest deviation ($d = 17.76$) occurred when $P = 72$, SAT = .5 and the sample size was 144. Finally, the Bartlett test showed its greatest deviation ($d = 6.84$) when P was 72, SAT = .8 and unique variables were present. The

Table B-3

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables Per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation when P = 36, N = 72, with 6 Variables Per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	-1.60	-.40	-.10	7.00	-2.00	-3.20	-3.40
		(sd)	(1.14)	(0.89)	(1.52)	(0.0)	(1.00)	(0.84)	(1.14)
	Y	d	-3.00	0.80	0.60	8.00	-3.60	-4.40	-4.60
		(sd)	(0.71)	(1.30)	(1.08)	(1.22)	(1.14)	(0.89)	(0.89)
Y	UNIQ	N d	-1.00	0.0	0.80	6.80	-2.20	-3.20	-3.40
		(sd)	(0.71)	(0.71)	(1.30)	(0.84)	(1.64)	(1.30)	(1.14)
	Y	d	-2.60	-.40	0.10	8.20	-2.80	-4.60	-4.60
		(sd)	(1.95)	(1.14)	(0.96)	(0.84)	(2.17)	(0.89)	(0.89)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.10	0.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(0.0)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.50	3.60	3.40	2.40	1.80
		(sd)	(0.0)	(0.0)	(0.50)	(0.55)	(0.89)	(0.89)	(0.45)
Y	UNIQ	N d	0.0	0.0	0.40	0.0	0.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.55)	(0.0)	(0.45)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.40	3.60	4.20	2.40	2.40
		(sd)	(0.0)	(0.0)	(0.89)	(0.55)	(0.84)	(0.55)	(0.55)

Table B-4

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation when P = 36, N = 72, with 6 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	-1.80	-.80	1.60	7.20	-2.20	-3.40	-3.60
		(sd)	(0.84)	(0.84)	(0.55)	(0.84)	(0.84)	(0.89)	(0.55)
	Y	d	-3.20	0.40	0.70	8.00	-2.20	-4.00	-4.00
		(sd)	(2.17)	(1.14)	(1.92)	(0.71)	(1.10)	(1.22)	(1.22)
	UNIQ	N d	-1.20	-.40	1.10	6.80	-2.00	-3.80	-3.80
		(sd)	(0.84)	(0.89)	(0.55)	(0.45)	(1.00)	(0.84)	(0.84)
Y	Y	d	-3.60	-.20	0.40	7.80	-3.00	-4.80	-4.80
		(sd)	(2.07)	(1.10)	(1.56)	(0.45)	(0.71)	(0.45)	(0.45)
Saturation = .8									
N	UNIQ	N d	0.20	0.0	0.10	0.0	0.60	0.0	0.0
		(sd)	(0.45)	(0.0)	(0.22)	(0.0)	(1.34)	(0.0)	(0.0)
	Y	d	0.20	-.20	0.20	3.40	3.80	2.60	2.40
		(sd)	(0.45)	(0.45)	(0.45)	(0.54)	(0.84)	(0.89)	(0.45)
	UNIQ	N d	0.60	0.0	0.20	0.40	0.20	0.0	0.0
		(sd)	(0.55)	(0.0)	(0.45)	(0.54)	(0.45)	(0.0)	(0.0)
Y	Y	d	0.20	0.0	0.70	3.20	5.40	3.40	3.20
		(sd)	(0.45)	(0.0)	(0.84)	(0.84)	(2.41)	(1.95)	(1.64)

Table B-5

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation when P = 36, N = 180, with 12 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	0.0	0.0	0.70	7.20	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(1.30)	(1.10)	(0.0)	(0.0)	(0.0)
	Y d	0.0	0.0	0.30	8.80	0.0	0.0	0.0	
		(sd)	(0.0)	(0.0)	(0.45)	(1.30)	(0.0)	(0.0)	(0.0)
Y	UNIQ	N d	0.0	0.0	0.50	7.40	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.71)	(0.89)	(0.0)	(0.0)	(0.0)
	Y d	0.0	0.0	1.20	8.20	0.20	0.0	0.0	
		(sd)	(0.0)	(0.0)	(1.15)	(0.84)	(0.45)	(0.0)	(0.0)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
	Y d	0.0	0.0	0.0	2.40	3.0	2.20	2.20	
		(sd)	(0.0)	(0.0)	(0.0)	(0.55)	(0.0)	(0.45)	(0.45)
Y	UNIQ	N d	0.0	0.0	0.0	0.0	0.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.45)	(0.0)	(0.0)
	Y d	0.0	0.0	0.10	1.80	2.80	2.40	2.20	
		(sd)	(0.0)	(0.0)	(0.22)	(0.45)	(0.45)	(0.55)	(0.45)

Table B-6

Means and Standard Deviations of the Difference from the MJC with and without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 36, N = 180, with 12 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	0.0	0.0	0.80	7.80	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.84)	(0.45)	(0.0)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.20	8.60	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.45)	(0.55)	(0.0)	(0.0)	(0.0)
	UNIQ	N d	0.0	0.0	0.20	7.40	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.45)	(0.55)	(0.0)	(0.0)	(0.0)
Y	Y	d	0.0	0.0	0.60	8.20	0.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.42)	(0.45)	(0.45)	(0.0)	(0.0)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.40	0.0	1.60	0.20	0.20
		(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(1.14)	(0.45)	(0.45)
	Y	d	0.0	0.0	0.60	2.40	3.60	2.80	2.80
		(sd)	(0.0)	(0.0)	(0.82)	(0.55)	(0.54)	(0.45)	(0.45)
	UNIQ	N d	0.0	0.0	0.10	0.0	0.60	0.20	0.0
		(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(0.55)	(0.45)	(0.0)
Y	Y	d	0.0	0.0	0.0	2.0	4.20	3.20	3.20
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.45)	(0.45)	(0.45)

Table B-7

Means and Standard Deviations of the Difference from the MJC with and without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 36, N = 180, with 6 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
UNIQ N	d		-1.60	0.0	0.70	6.40	-.80	-1.40	-1.60
		(sd)	(0.89)	(0.0)	(0.45)	(0.55)	(0.45)	(0.55)	(0.55)
	Y d	-2.80	0.0	-.30	7.20	-1.00	-1.80	-2.00	
		(sd)	(0.84)	(0.71)	(0.67)	(0.45)	(0.71)	(0.84)	(0.71)
UNIQ N	d		-.40	0.0	0.10	5.60	-.60	-.80	-.80
		(sd)	(0.55)	(0.0)	(1.14)	(1.14)	(0.55)	(0.84)	(0.84)
	Y d	-3.80	0.40	0.20	7.60	-.60	-1.80	-1.80	
		(sd)	(1.79)	(0.55)	(0.84)	(0.55)	(0.89)	(0.45)	(0.45)
Saturation = .8									
UNIQ N	d		0.0	0.0	0.0	0.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
	Y d	0.0	0.0	0.10	3.80	5.60	5.00	5.00	
		(sd)	(0.0)	(0.0)	(0.22)	(0.45)	(0.55)	(0.71)	(0.71)
UNIQ N	d		0.0	0.0	0.10	0.0	0.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(0.45)	(0.0)	(0.0)
	Y d	0.0	0.0	0.70	3.20	5.80	5.20	5.20	
		(sd)	(0.0)	(0.0)	(1.04)	(0.84)	(0.45)	(0.45)	(0.45)

Table B-8

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 36, N = 180, with 6 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N d	-1.40	0.0	1.40	5.80	-.80	-1.0	-1.40
		(sd)	(0.55)	(0.0)	(0.65)	(0.84)	(0.45)	(0.0)	(0.54)
N	Y	d	-2.80	0.20	1.60	7.40	-.80	-1.80	-2.00
		(sd)	(1.48)	(0.45)	(1.08)	(0.89)	(0.84)	(0.84)	(1.00)
	UNIQ	N d	-.80	0.0	1.50	6.20	-.40	-1.00	-1.20
		(sd)	(1.30)	(0.0)	(0.87)	(0.45)	(0.55)	(0.0)	(0.45)
Y	Y	d	-4.60	1.00	0.70	7.40	-.60	-1.60	-1.80
		(sd)	(1.95)	(0.71)	(1.20)	(0.89)	(0.89)	(1.14)	(0.89)
Saturation = .8									
	UNIQ	N d	0.0	0.0	-.50	0.0	0.80	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.71)	(0.0)	(0.84)	(0.0)	(0.0)
N	Y	d	0.0	0.0	0.60	3.40	6.40	5.80	5.80
		(sd)	(0.0)	(0.0)	(0.42)	(0.55)	(0.55)	(0.45)	(0.45)
	UNIQ	N d	0.0	0.0	0.50	0.0	1.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.87)	(0.0)	(0.45)	(0.0)	(0.0)
Y	Y	d	0.0	0.0	0.10	3.00	7.80	6.60	6.20
		(sd)	(0.0)	(0.0)	(0.22)	(0.71)	(0.84)	(0.55)	(0.45)

Table B-10

Means and Standard Deviations of the Difference from the MJC with and without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 72, N = 144, with 12 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	0.0	0.0	0.90	18.00	0.80	-.40	-.40
		(sd)	(0.0)	(0.0)	(0.55)	(1.22)	(1.10)	(0.55)	(0.55)
	Y	d	0.0	0.0	0.20	18.80	-.20	-1.00	-1.20
		(sd)	(0.0)	(0.0)	(0.45)	(0.84)	(0.45)	(0.71)	(0.45)
Y	UNIQ	N d	0.0	0.0	0.90	19.00	-.20	-.80	-.80
		(sd)	(0.0)	(0.0)	(1.52)	(0.71)	(0.45)	(0.45)	(0.45)
	Y	d	0.0	0.0	0.70	19.00	-.20	-.80	-1.00
		(sd)	(0.0)	(0.0)	(1.30)	(0.71)	(0.84)	(0.84)	(0.71)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.0	0.20	2.00	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.45)	(0.71)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.0	4.80	6.00	4.40	4.00
		(sd)	(0.0)	(0.0)	(0.0)	(0.45)	(3.67)	(2.79)	(3.08)
Y	UNIQ	N d	0.0	0.0	0.40	1.00	1.80	0.40	0.20
		(sd)	(0.0)	(0.0)	(0.65)	(0.71)	(1.64)	(0.54)	(0.45)
	Y	d	0.0	0.0	0.50	5.20	6.40	4.60	4.60
		(sd)	(0.0)	(0.0)	(0.50)	(0.45)	(0.89)	(1.14)	(1.14)

Table B-11

Means and Standard Deviations of the Difference from the MJC with and without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and $P = 72$, $N = 144$, with 8 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N	d	-1.80	-.80	0.20	15.80	-1.20	-2.80
			(sd)	(1.30)	(0.45)	(0.91)	(0.45)	(1.48)	(0.84)
N	Y	d		-2.20	-.60	0.80	17.20	-1.80	-3.40
			(sd)	(0.45)	(0.55)	(2.02)	(0.84)	(0.84)	(0.55)
	UNIQ	N	d	-.40	0.0	1.60	16.60	-.40	-1.60
			(sd)	(0.55)	(0.0)	(0.55)	(0.55)	(1.34)	(0.89)
Y	Y	d		-.40	0.80	2.70	17.20	-1.20	-3.0
			(sd)	(0.89)	(0.84)	(1.79)	(1.30)	(0.84)	(0.45)
Saturation = .8									
	UNIQ	N	d	0.0	0.0	0.0	0.0	0.0	0.0
			(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
N	Y	d		0.0	0.0	0.60	6.20	7.40	5.80
			(sd)	(0.0)	(0.0)	(0.65)	(0.45)	(1.52)	(1.10)
	UNIQ	N	d	0.0	0.0	0.50	0.0	0.40	0.0
			(sd)	(0.0)	(0.0)	(0.50)	(0.0)	(0.55)	(0.0)
Y	Y	d		0.0	0.0	1.00	6.40	7.60	5.60
			(sd)	(0.0)	(0.0)	(0.94)	(0.89)	(1.67)	(1.14)

Table B-12

Means and Standard Deviations of the Difference from the MJC with and without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 72, N = 144, with 8 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N d	-1.00	-.20	0.80	16.00	0.0	-2.00	-2.20
		(sd)	(1.00)	(0.84)	(1.92)	(1.00)	(1.58)	(1.00)	(0.84)
N	Y	d	-2.60	-.60	0.70	18.0	-1.60	-3.60	-3.60
		(sd)	(0.55)	(0.89)	(1.30)	(0.0)	(0.55)	(0.55)	(0.55)
	UNIQ	N d	-.20	0.20	1.70	16.20	0.20	-1.40	-1.40
		(sd)	(0.45)	(0.45)	(1.10)	(0.45)	(0.84)	(0.55)	(0.55)
Y	Y	d	-.40	0.40	1.70	17.40	-1.00	-2.60	-2.80
		(sd)	(0.89)	(0.55)	(1.20)	(0.55)	(0.71)	(0.55)	(0.45)
Saturation = .8									
	UNIQ	N d	0.0	0.0	0.0	0.20	1.00	0.20	0.20
		(sd)	(0.0)	(0.0)	(0.0)	(0.45)	(1.73)	(0.45)	(0.45)
N	Y	d	0.0	0.0	0.50	6.40	9.20	7.00	6.80
		(sd)	(0.0)	(0.0)	(0.50)	(0.55)	(2.39)	(1.58)	(1.64)
	UNIQ	N d	0.20	0.0	0.30	0.40	2.00	0.0	0.0
		(sd)	(0.45)	(0.0)	(0.45)	(0.55)	(1.00)	(0.0)	(0.0)
Y	Y	d	0.0	0.0	0.10	6.20	8.40	6.80	6.20
		(sd)	(0.0)	(0.0)	(0.22)	(0.45)	(1.14)	(0.84)	(0.84)

Table B-13

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and $P = 72$, $N = 380$, with 12 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	0.0	0.0	0.60	15.00	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.89)	(0.71)	(0.0)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.10	16.40	0.80	0.20	0.20
		(sd)	(0.0)	(0.0)	(0.22)	(0.55)	(1.30)	(0.45)	(0.45)
	UNIQ	N d	0.0	0.0	0.30	15.00	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.45)	(0.0)	(0.0)	(0.0)	(0.0)
Y	Y	d	0.0	0.0	0.0	16.80	0.40	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(1.10)	(0.89)	(0.0)	(0.0)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.0	0.0	0.20	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.45)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.10	4.20	5.80	5.20	5.20
		(sd)	(0.0)	(0.0)	(0.22)	(0.45)	(0.45)	(0.45)	(0.45)
	UNIQ	N d	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Y	Y	d	0.0	0.0	3.0	4.40	5.80	5.20	5.00
		(sd)	(0.0)	(0.0)	(0.0)	(0.55)	(0.45)	(0.84)	(0.71)

Table B-14

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 72, N = 380, with 12 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	0.0	0.0	0.60	15.00	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.55)	(1.22)	(0.0)	(0.0)	(0.0)
	Y d	0.0	0.0	0.0	17.0	0.80	0.0	0.0	
		(sd)	(0.0)	(0.0)	(0.0)	(1.0)	(0.84)	(0.0)	(0.0)
	UNIQ	N d	0.0	0.0	0.90	15.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.89)	(1.0)	(0.0)	(0.0)	(0.0)
Y	Y d	0.0	0.0	0.10	16.60	0.40	0.0	0.0	
		(sd)	(0.0)	(0.0)	(0.74)	(0.55)	(0.55)	(0.0)	(0.0)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.0	0.0	2.20	0.80	0.60
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.84)	(0.45)	(0.55)
	Y d	0.0	0.0	0.20	4.60	7.20	6.20	6.20	
		(sd)	(0.0)	(0.0)	(0.27)	(0.55)	(1.10)	(0.45)	(0.45)
	UNIQ	N d	0.0	0.0	0.0	0.0	3.80	1.80	1.80
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.45)	(0.45)	(0.45)
Y	Y d	0.0	0.0	0.40	4.40	9.00	7.40	7.0	
		(sd)	(0.0)	(0.0)	(0.65)	(0.55)	(1.87)	(1.34)	(1.0)

Table B-15

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 72, N = 380, with 8 Variables per Component and Complex Variables Absent.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
N	UNIQ	N d	-2.00	0.0	0.0	14.40	0.0	-.80	-1.0
		(sd)	(0.71)	(0.0)	(0.0)	(0.55)	(0.0)	(0.45)	(0.0)
	Y	d	-2.20	-.20	0.90	15.40	-.20	-1.0	-1.20
		(sd)	(0.45)	(0.45)	(0.74)	(0.89)	(0.45)	(0.0)	(0.45)
	UNIQ	N d	0.0	0.0	0.20	13.80	0.0	-.40	-.60
		(sd)	(0.0)	(0.0)	(0.45)	(0.45)	(0.0)	(0.55)	(0.55)
	Y	d	0.0	0.20	0.60	15.80	0.60	0.0	-.20
		(sd)	(0.0)	(0.45)	(0.89)	(0.84)	(0.89)	(0.0)	(0.45)
Saturation = .8									
N	UNIQ	N d	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.20	5.80	8.40	7.80	7.60
		(sd)	(0.0)	(0.0)	(0.27)	(0.84)	(0.55)	(0.45)	(0.55)
	UNIQ	N d	0.0	0.0	0.0	0.0	0.40	0.0	0.0
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(0.89)	(0.0)	(0.0)
	Y	d	0.0	0.0	0.20	5.80	8.60	7.80	7.60
		(sd)	(0.0)	(0.0)	(0.45)	(0.45)	(0.55)	(0.84)	(0.55)

Table B-16

Means and Standard Deviations of the Difference from the MJC with or without Equal Number of Variables per Component (EP/m = N, Y) and Unique Variables (UNIQ = N, Y) at Low and High Component Saturation and P = 72, N = 380, with 8 Variables per Component and Complex Variables Present.

			Method						
EP/m			MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5									
	UNIQ	N d	-0.60	0.0	1.10	13.80	0.20	-.60	-0.60
		(sd)	(0.55)	(0.0)	(0.89)	(0.84)	(0.45)	(0.55)	(0.55)
N	Y	d	-3.20	0.20	0.50	16.20	1.20	-0.20	-0.20
		(sd)	(1.10)	(0.45)	(1.50)	(0.84)	(1.30)	(0.84)	(0.84)
	UNIQ	N d	0.0	0.0	1.20	13.80	0.20	-.20	-.40
		(sd)	(0.0)	(0.0)	(0.84)	(0.84)	(0.45)	(0.45)	(0.55)
Y	Y	d	0.0	0.40	1.10	15.60	1.20	0.0	-.20
		(sd)	0.0	0.55	0.65	0.55	0.84	0.71	0.45)
Saturation = .8									
	UNIQ	N d	0.0	0.0	0.10	0.0	2.60	0.80	0.60
		(sd)	(0.0)	(0.0)	(0.22)	(0.0)	(0.55)	(0.45)	(0.55)
N	Y	d	0.0	0.0	0.70	5.20	11.00	9.80	9.60
		(sd)	(0.0)	(0.0)	(1.86)	(0.45)	(1.22)	(1.30)	(0.89)
	UNIQ	N d	0.0	0.0	0.0	0.0	2.80	1.40	1.20
		(sd)	(0.0)	(0.0)	(0.0)	(0.0)	(1.30)	(0.89)	(0.84)
Y	Y	d	0.0	0.0	0.60	5.80	11.40	9.60	9.40
		(sd)	(0.0)	(0.0)	(0.55)	(0.45)	(1.14)	(0.89)	(0.89)

worst cases for 4 of the methods occurred when the saturation was low. The Bartlett test was the only exception to the pattern.

Tables B-27 and B-28 present the percent of each decision method's estimates that deviated a set amount from the population MJC. These tables provide an overview of the methods' performance across all the conditions presented in this study. An examination of the proportion of each method's decision that fall ± 1 from the criterion across all cases examined indicates the PA method was the most accurate (98.4% ± 1 from the population MJC). The MAP and SCREE methods followed with overall percentages of 87.6 and 83.4 respectively. The Bartlett test's percentages ranged from 59.9 to 61.4 for alpha levels of .0005 and .05 respectively. The K1 rule came within one of the criterion in only 50.0 percent of the cases examined.

Table B-17

Means and Standard Deviations of the Difference from the Population

Value of MJC at each Level of Sample Size and Saturation when P = 36.

Sample Size		Method						
		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
72	d	-1.11	-0.04	0.64	8.36	-1.99	-2.41	-2.49
	(sd)	(1.61)	(0.75)	(1.06)	(1.21)	(1.64)	(1.72)	(1.73)
180	d	-1.14	0.10	0.65	7.32	-0.50	-0.70	-0.79
	(sd)	(1.71)	(0.38)	(0.93)	(1.16)	(0.69)	(0.88)	(0.95)
Saturation = .8								
72	d	0.09	-0.01	0.26	1.74	1.44	0.98	0.85
	(sd)	(0.28)	(0.11)	(0.48)	(1.50)	(1.69)	(1.32)	(1.21)
180	d	0.0	0.0	0.18	1.38	2.40	2.10	2.05
	(sd)	(0.0)	(0.0)	(0.52)	(1.50)	(2.48)	(2.38)	(2.34)

Table B-18

Means and Standard Deviations of the Difference from the Population Value
of MJC when Complex Variables are Absent or Present, Saturation is Low
or High and when P = 36.

		Method						
Complex Variables		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
NO	d	-1.04	0.02	0.41	7.80	-1.32	-1.54	-1.62
	(sd)	(1.51)	(0.57)	(0.94)	(1.32)	(1.56)	(1.61)	(1.66)
YES	d	-1.21	0.38	0.88	7.89	-1.16	-1.58	-1.65
	(sd)	(1.79)	(0.62)	(1.00)	(1.27)	(1.36)	(1.61)	(1.62)
Saturation = .8								
NO	d	0.0	0.0	0.18	1.58	1.65	1.34	1.25
	(sd)	(0.0)	(0.0)	(0.44)	(1.58)	(1.96)	(1.78)	(1.76)
YES	d	0.09	0.01	0.26	1.54	2.19	1.74	1.65
	(sd)	(0.28)	(0.11)	(0.55)	(1.54)	(2.39)	(2.18)	(2.12)

Table B-19

Means and Standard Deviations of the Difference from the Population Value of MJC when there are Equal or Unequal Numbers of Variables per Component (EP/m = N, Y), Saturation is Low or High and when P = 36.

		Method						
EP/m		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
N	d	-1.14	0.02	0.74	7.90	-1.30	-1.55	-1.66
	(sd)	(1.46)	(0.62)	(1.04)	(1.29)	(1.42)	(1.52)	(1.58)
Y	d	-1.11	0.38	0.55	7.79	-1.19	-1.56	-1.61
	(sd)	(1.84)	(0.58)	(0.94)	(1.30)	(1.51)	(1.70)	(1.69)
Saturation = .8								
N	d	0.02	-0.01	0.18	1.60	1.81	1.42	1.36
	(sd)	(0.16)	(0.11)	(0.45)	(1.62)	(2.00)	(1.89)	(1.86)
Y	d	0.06	0.0	0.26	1.51	2.02	1.65	1.54
	(sd)	(0.24)	(0.0)	(0.54)	(1.50)	(2.33)	(2.11)	(2.05)

Table B-20

Means and Standard Deviations of the Difference from the Population Value
of MJC when Unique Variables are Absent or Present (UNIQ = N, Y),
Saturation is Low or High and when P = 36.

		Method						
UNIQ		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
N	d	-0.60	-0.10	0.74	7.35	-1.04	-1.32	-1.42
	(sd)	(0.91)	(0.44)	(0.97)	(1.31)	(1.24)	(1.38)	(1.44)
Y	d	-1.65	0.16	0.55	8.34	-1.45	-1.79	-1.85
	(sd)	(2.03)	(0.70)	(1.01)	(1.07)	(1.64)	(1.78)	(1.79)
Saturation = .8								
N	d	0.06	0.0	0.14	0.11	0.21	0.04	0.01
	(sd)	(0.24)	(0.0)	(0.44)	(0.32)	(0.47)	(0.19)	(0.11)
Y	d	0.25	0.01	0.29	3.00	3.62	3.04	2.89
	(sd)	(0.16)	(0.11)	(0.54)	(0.75)	(1.83)	(1.86)	(1.86)

Table B-21

Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of the Number of Variables per Component (#P/MJC = 6, 12) and Saturation when P = 36.

		Method						
#P/MJC		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
6	d	-2.26	0.04	0.69	7.09	-2.18	-2.66	-2.80
	(sd)	(1.70)	(0.83)	(1.15)	(1.02)	(1.52)	(1.56)	(1.51)
12	d	0.01	0.02	0.59	8.60	-0.31	-0.45	-0.48
	(sd)	(0.11)	(0.16)	(0.81)	(1.07)	(0.49)	(0.55)	(0.57)
Saturation = .8								
6	d	0.08	-0.01	0.26	1.72	2.45	2.09	2.00
	(sd)	(0.26)	(0.11)	(0.60)	(1.75)	(2.63)	(2.45)	(2.37)
12	d	0.01	0.0	0.18	1.39	1.39	0.99	0.90
	(sd)	(0.11)	(0.0)	(0.37)	(1.33)	(1.40)	(1.20)	(1.20)

Table B-22

Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of Sample Size and Saturation when $P = 72$.

Sample Size		Method						
		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
144	d	-0.54	-0.04	1.08	17.76	-0.98	-1.61	-1.76
	(sd)	(1.02)	(0.56)	(1.33)	(1.34)	(1.17)	(1.27)	(1.17)
380	d	-0.50	0.04	0.51	15.35	0.02	-0.19	-0.26
	(sd)	(1.04)	(0.25)	(0.78)	(1.26)	(0.55)	(0.48)	(0.52)
Saturation = .8								
144	d	0.01	0.0	0.26	2.96	3.08	2.51	2.40
	(sd)	(0.11)	(0.0)	(0.52)	(2.81)	(3.16)	(2.84)	(2.74)
380	d	0.0	0.0	0.26	2.51	4.49	3.99	3.86
	(sd)	(0.0)	(0.0)	(0.90)	(2.60)	(3.83)	(3.69)	(3.62)

Table B-23

Means and Standard Deviations of the Difference from the Population Value of MJC at each Level of the Number of Variables per Component (#P/MJC = 8, 12) and Saturation when $P = 72$.

P/MJC		Method						
		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
8	d	-1.06	-0.01	0.99	15.82	-0.86	-1.46	-1.58
	(sd)	(1.22)	(0.60)	(1.26)	(1.46)	(1.22)	(1.34)	(1.30)
12	d	0.02	0.01	0.60	17.29	-0.09	-0.34	-0.45
	(sd)	(0.16)	(0.11)	(0.94)	(1.77)	(0.62)	(0.65)	(0.67)
Saturation = .8								
8	d	0.01	0.0	0.21	3.02	4.48	3.91	3.76
	(sd)	(0.11)	(0.0)	(0.67)	(3.01)	(4.07)	(3.86)	(3.77)
12	d	0.0	0.0	0.30	2.45	3.09	2.59	2.50
	(sd)	(0.0)	(0.0)	(0.79)	(2.34)	(2.85)	(2.64)	(2.59)

Table B-24

Means and Standard Deviations of the Difference from the Population Value of MJC when Complex Variables are Absent or Present, Saturation is Low or High and when $P = 72$.

		Method						
Complex Variables		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
N	d	-0.54	-0.02	0.77	16.52	-0.54	-0.95	-1.10
	(sd)	(1.00)	(0.44)	(1.17)	(1.74)	(1.08)	(1.23)	(1.20)
Y	d	-0.50	0.02	0.82	16.59	-0.41	-0.85	-0.92
	(sd)	(1.06)	(0.42)	(1.08)	(1.82)	(1.00)	(1.16)	(1.16)
Saturation = .8								
N	d	0.0	0.0	0.36	2.70	2.98	2.68	2.61
	(sd)	(0.0)	(0.0)	(0.81)	(2.77)	(3.24)	(3.03)	(2.94)
Y	d	0.01	0.0	0.15	2.78	4.59	3.82	3.65
	(sd)	(0.11)	(0.0)	(0.64)	(2.65)	(3.72)	(3.60)	(3.53)

Table B-25

Means and Standard Deviations of the Difference from the Population Value of MJC when there are Equal or Unequal Numbers of Variables per Component (EP/m = N, Y), Saturation is Low or High and when P = 72.

		Method						
EP/m		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
N	d	-0.98	-0.14	0.56	16.55	-0.64	-1.05	-1.15
	(sd)	(1.24)	(0.44)	(1.00)	(1.78)	(1.24)	(1.35)	(1.31)
Y	d	-0.06	0.14	1.02	16.56	-0.31	-0.75	-0.88
	(sd)	(0.40)	(0.38)	(1.19)	(1.78)	(0.76)	(1.00)	(1.01)
Saturation = .8								
N	d	0.0	0.0	0.08	2.69	3.68	3.18	3.06
	(sd)	(0.0)	(0.0)	(0.58)	(2.71)	(3.61)	(3.41)	(3.36)
Y	d	0.01	0.0	0.43	2.79	3.89	3.32	3.20
	(sd)	(0.11)	(0.0)	(0.82)	(2.72)	(3.54)	(3.34)	(3.23)

Table 8-26

Means and Standard Deviations of the Difference from the Population Value of MJC when Unique Variables are Absent or Present (UNIQ = N, Y), Saturation is Low or High and when P = 72.

		Method						
UNIQ		MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5								
N	d	-0.35	-0.05	0.81	15.89	-0.35	-0.74	-0.82
	(sd)	(0.81)	(0.31)	(1.02)	(1.85)	(0.86)	(0.99)	(0.95)
Y	d	-0.64	0.05	0.78	17.22	-0.60	-1.06	-1.20
	(sd)	(1.18)	(0.52)	(1.22)	(1.42)	(1.19)	(1.35)	(1.34)
Saturation = .8								
N	d	0.01	0.0	0.08	0.12	0.72	0.34	0.29
	(sd)	(0.11)	(0.0)	(0.27)	(0.37)	(1.04)	(0.64)	(0.50)
Y	d	0.0	0.0	0.43	5.35	6.84	6.16	5.98
	(sd)	(0.0)	(0.0)	(0.97)	(0.87)	(2.28)	(2.28)	(2.23)

Table B-27

Percent of Decisions at Different Levels of Deviation from the
Population Value at Low and High Saturation when P = 36.

	Method						
	MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5							
+3	0.0	0.6	4.4	100.0	0.0	0.0	0.0
+2	0.0	1.2	19.4	0.0	0.0	0.0	0.0
+1	0.6	8.7	32.5	0.0	0.0	0.0	0.0
+0	56.9	82.5	35.0	0.0	40.6	30.6	29.4
-1	11.9	3.7	6.8	0.0	29.4	33.1	30.6
-2	11.2	3.1	1.8	0.0	11.2	12.5	14.4
-3	19.3	0.0	0.0	0.0	18.8	23.8	25.6
Saturation = .8							
+3	0.0	0.0	0.6	37.6	39.4	26.2	23.7
+2	0.0	0.0	3.8	11.9	6.3	15.0	16.9
+1	4.4	0.0	22.5	6.3	12.5	7.5	5.0
+0	95.6	99.4	71.9	44.4	41.9	51.2	54.4
-1	0.0	0.6	0.6	0.0	0.0	0.0	0.0
-2	0.0	0.0	0.6	0.0	0.0	0.0	0.0
-3	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table B-28

Percent of Decisions at Different Levels of Deviation from the
Population Value at Low and High Saturation when P = 72.

	Method						
	MAP	PA	SCREE	K1	BA	BB	BC
Saturation = .5							
+3	0.0	0.0	10.0	100.0	0.6	0.0	0.0
+2	0.0	0.6	16.2	0.0	1.2	0.6	0.0
+1	1.9	6.9	25.0	0.0	5.6	1.9	1.9
±0	71.5	85.0	43.8	0.0	56.3	45.6	40.0
-1	7.5	6.9	4.3	0.0	21.2	28.1	31.9
-2	11.2	0.6	0.6	0.0	9.4	10.0	11.2
-3	7.5	0.0	0.0	0.0	5.6	13.7	14.9
Saturation = .8							
+3	0.0	0.0	3.7	50.0	55.0	47.5	46.2
+2	0.0	0.0	1.9	0.6	6.9	6.9	7.5
+1	0.6	0.0	21.3	5.0	7.5	8.1	6.9
±0	99.4	100.0	71.9	44.4	30.6	37.5	39.4
-1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-2	0.0	0.0	0.6	0.0	0.0	0.0	0.0
-3	0.0	0.0	0.6	0.0	0.0	0.0	0.0

APPENDIX C
CONSTRUCTION OF POPULATION MATRICES

The population correlation matrices for this study were constructed from the general model

$$R = AA' + D^2 \quad (C-1)$$

The steps to create the matrix were as follows:

- (1) An appropriate population component pattern (A) was created in accordance with the level of the number of variables (P), the level of the saturation factor (SAT), the level of the number of components factor (m), the presence or absence of equal numbers of variables per component, unique and complex variables.
- 2) This matrix (A) was post-multiplied by its transpose (AA') to create a matrix R*.
- 3) R* was augmented by replacing the diagonal elements with 1.0's. The introduction of one's into the diagonal produced a correlation of full rank and permitted subsequent analysis.

For example, in a 10 variable (P = 10), two component (m = 2), high saturation (SAT = .8), unequal variables per MJC, CV and UNIQ MNC present case (Complexity = 6), the underlying component structure A would have been:

$$A = \begin{bmatrix} .8 & 0 & 0 & 0 & .2 & 0 & 0 \\ .8 & 0 & 0 & 0 & 0 & .2 & 0 \\ .8 & 0 & 0 & 0 & .0 & 0 & .2 \\ .8 & 0 & 0 & 0 & .2 & 0 & 0 \\ .8 & 0 & 0 & 0 & 0 & .2 & 0 \\ 0 & 0 & 1.0 & 0 & 0 & 0 & 0 \\ 0 & .8 & 0 & 0 & .2 & 0 & 0 \\ 0 & .8 & 0 & 0 & 0 & .2 & 0 \\ 0 & .8 & 0 & 0 & 0 & 0 & .2 \\ 0 & 0 & 0 & 1.0 & 0 & 0 & 0. \end{bmatrix}$$

Given this component structure, R^* ($R^* = AA'$) would have been:

	.68	.64	.64	.68	.64	.00	.04	.00	.00	.00
	--	.68	.64	.64	.68	.00	.00	.04	.00	.00
	--	--	.68	.64	.68	.00	.00	.00	.04	.00
	--	--	--	.68	.64	.00	.04	.00	.00	.00
$R^* =$	--	--	--	--	.68	.00	.00	.04	.00	.00
	--	--	--	--	--	1.00	.00	.00	.00	.00
	--	--	--	--	--	--	.68	.64	.64	.00
	--	--	--	--	--	--	--	.68	.64	.00
	--	--	--	--	--	--	--	--	.68	.00
	--	--	--	--	--	--	--	--	--	1.00

The diagonal elements of D^2 would have been .32, .32, .32, .32, .32, .00, .32 .32 .32, .00. The off diagonal elements of D^2 would all be .00.

The resultant population R matrix ($R = R^* + D^2$) would equal the R^* matrix with 1.0's in the diagonal. Montaneilli's (1975) program was then employed to generate five sample correlation matrices, at each of two levels of N, from the population matrix.

BIBLIOGRAPHY

- Anderson, A. D., Acito, F. & Lee, H. (1982). A simulation study of three methods for determining the number of image components. Multivariate Behavioral Research, 17, 493-502.
- Anderson, T. W. & Rubin, H. (1956). Statistical inference in factor analysis. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5.
- Armstrong, J. S. & Soelberg, P. (1968). On the interpretation of factor analysis. Psychological Bulletin, 70, 361-364.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. The British Journal of Psychology, 3, 77-85.
- Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. The British Journal of Psychology, 4, 1-2.
- Browne, M. W. (1968). A note on lower bounds for the number of common factors. Psychometrika, 33, 233.
- Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.
- Cattell, R. B. & Jaspers, J. (1967). A general plasmode for factor analytic exercises and research. Multivariate Behavioral Research Monographs, 3, 1-212.
- 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, 289-325.
- Cliff, N. (1970). The relation between sample and population characteristic vectors. Psychometrika, 35, 2, 163.

- Cliff, N. & Hamburger, C. D. (1967). The study of sampling errors in factor analysis by means of artificial experiments. Psychological Bulletin, 68, 430-445.
- Cliff, N. & Pennell, R. (1967). The influence of communality, factor strength and loading size on the sample characteristics of factor loadings. Psychometrika, 32, 309.
- Crawford, C. B. (1975). Determining the number of interpretable factors. Psychological Bulletin, 82, 226-237.
- Crawford, C. B. & Koopman, P. (1979). Note: Interrater reliability of scree test and mean square ratio test of number of factors. Perceptual and Motor Skills, 49, 223-226.
- Dziuban, C. D. & Shirkey, E. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological Bulletin, 81, (6), 358-361.
- Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. Multivariate Behavioral Research, 18, 197-218.
- Glass, G. V. & Taylor, P. A. (1966). Factor analytic methodology. Review of Educational Research, 36, 566-587.
- Gorsuch, R. L. (1974). Factor Analysis. Philadelphia: Saunders.
- Green, S. B. (1983). Identifiability of spurious factors using linear factor analysis with binary items. Applied Psychological Measurement, 7, (2), 139-147.
- Guttman, L. (1954). Some necessary conditions for common factor analysis. Psychometrika, 19, 149-162.
- Hakstian, A. R., Rogers, W. T. & Cattell, R. B. (1982). The behavior of number of factors rules with simulated data. Multivariate Behavioral Research, 17, 193-219.

- Harris, M. L. & Harris, C. W. (1971). A factor analytic interpretation strategy. Educational and Psychological Measurement, 31, (3), 589.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185.
- Horn, J. L. (1969). On the internal consistency reliability of factors. Multivariate Behavioral Research, 4, 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, 283-300.
- Horst, P. (1965). Factor Analysis of Data Matrices. New York: Holt, Rinehart and Winston.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24, 417-441, 498-520.
- Humphreys, L. G. (1964). Number of cases and number of factors: An example where N is very large. Educational and Psychological Measurement, 24, 457.
- Humphreys, L. G. & Montanelli, R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. Multivariate Behavioral Research, 10, 193-206.
- Jackson, D. N. & Chan, D. W. (1980). Maximum-Likelihood estimation in common factor analysis: A cautionary note. Psychological Bulletin, 88, 502-508.
- Joreskog, K. G. (1962). On the statistical treatment of residuals in factor analysis. Psychometrika, 27, 335-354.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

- Kaiser, H. F. (1961). A note on Guttman's lower bound for the number of common factors. British Journal of Statistical Psychology, 14, 1-2.
- Kaiser, H. F. (1970). A second-generation little jiffy. Psychometrika, 35, 401-415.
- Kaiser, H. F. & Caffrey, J. (1965). Alpha factor analysis. Psychometrika, 30, 1-44.
- Kaiser, H. F. & Hunka, S. (1973). Some empirical results with Guttman's stronger lower bound for the number of common factors. Educational and Psychological Measurement, 33, 99-102.
- Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. Proceedings of the Royal Society of Edinburgh, 60, 64-82.
- Lee, H. B. & Comrey, A. L. (1979). Distortions in a commonly used factor analytic procedure. Multivariate Behavioral Research, 14, 301-321.
- Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. Psychometrika, 33, 37-71.
- Montanelli, R. G., Jr. (1975). A computer program to generate sample correlation and covariance matrices. Educational and Psychological Measurement, 35, 195-197.
- Montanelli, R. G. & Humphreys, L. G. (1976). Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. Psychometrika, 41, 341-348.
- Morrison, D. F. (1976). Multivariate Statistical Methods, 2nd ed. McGraw-Hill, Kogakuska Ltd.

- Mote, T. A. (1970). An artifact of the rotation of too few factors: Study orientation vs. trait anxiety. Revista Interamericana de Psicologia, 37, 61-91.
- 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.
- Schonemann, P. H. & Steiger, J. H. (1978). On the validity of indeterminant factor scores. Bulletin of the Psychonomic Society, 12, 287-290.
- Schonemann, P. H. & Wang, M. (1972). Some new results on factor indeterminacy. Psychometrika, 37, 61-91.
- Steiger, J. H. & Schonemann, P. H. (1979). A history of factor indeterminacy. In Shye, S. (Ed.) Theory Construction and Data Analysis in the Behavioral Sciences. San Francisco: Jossey-Bass.
- Tucker, L. B., Koopman, R. F. & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. Psychometrika, 34, 421.
- Van de Geer, J. P. (1971). Introduction to Multivariate Analysis for the Social Sciences. San Francisco: Freeman.
- Veldman, D. J. (1979). Simple structure and the number of factors problem. Multivariate Behavioral Research, 14, 191-200.
- Velicer, W. F. (1974). A comparison of the stability of factor analysis, principal component analysis and rescaled image analysis. Educational and Psychological Measurement, 34, 563-572.
- Velicer, W. F. (1976a). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 3, 321-327.

- Velicer, W. F. (1976b). The relation between factor score estimates, image scores and principal component scores. Educational and Psychological Measurement, 36, 149-159.
- Velicer, W. F. (1977). An empirical comparison of the similarity of principal component, image, and factor patterns. Multivariate Behavioral Research, 12, 3-22.
- 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, 371-388.
- 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.
- Zwick, W. R. & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. Multivariate Behavioral Research, 17, 253-269.