2009

Determining the number of factors in data containing a single outlier: a study of factor analysis of simulated data

Victor Snipes Swaim
Louisiana State University and Agricultural and Mechanical College, vswaim1@lsu.edu

DETERMINING THE NUMBER OF FACTORS IN DATA CONTAINING A SINGLE OUTLIER: A STUDY OF FACTOR ANALYSIS OF SIMULATED DATA

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Educational Theory,
Policy, and Practice

by
Victor Snipes Swaim
B.S., Middle Tennessee State University, 1991
M.S., Louisiana State University, 1997
December 2009

DEDICATION

To my mother, Olivia Jeanette Swaim.

And

The memory of my father, Harry Edward Swaim Sr..

TABLE OF CONTENTS

ABSTRACT

Numerous procedures have been suggested for determining the number of factors to retain in factor analysis. However, previous studies have focused on comparing methods using normal data sets. This study had two phases. The first phase explored the Kaiser method, Scree test, Bartlett's chi-square test, Minimum Average Partial (1976 & 2000), Horn's parallel analysis, and Longman's Parallel Analysis on normal data using the estimation methods of Maximum Likelihood (ML), Principal Component Analysis (PCA), and Principal Factor Analysis (PFA). The second phase explored the Kaiser method, Scree test, Minimum Average Partial (1976 & 2000), and Horn's parallel analysis, and Longman's Parallel Analysis on data that contained outliers using the estimation methods of PCA and PFA. In the first phase, sample correlation matrices were generated with varied conditions (sample size, number of variables, estimation methods). Three hundred sample correlation matrices were generated for each condition for a grand total of eighteen hundred. The performance of parallel analysis and the Kaiser method were generally the best across all situations. However, the increase in variables and sample size under each condition showed a difference in accuracy among the methods. The increase in sample size resulted in little difference between estimation methods of PCA and PFA. Recommendations concerning the accuracy of the methods under each condition are discussed. In the second phase, fifty sample correlation matrices were randomly selected from each of the three hundred sample correlations matrices under each condition. An outlier was randomly incorporated in each of the fifty sample correlation matrices. The squared Mahalanobis distance was recorded for each to determine the distance at which the methods start to fail. The research conducted here indicates that Parallel Analysis and

Longman's Parallel Analysis was very resistant to outliers in some specific cases. However, it was evident from the data that each method tended to make the incorrect decision on retaining the correct number of factors when the squared Mahalanobis distance reached a certain amount. A discussion of method performance is given on each of the conditions to help determine the most effective and useful combinations on dealing with the outliers.

CHAPTER 1: INTRODUCTION

Statement of Problem

Factor analysis is used in research studies across a wide range of disciplines. For example, in the field of Psychology, factor analysis is commonly associated with intelligence research. The field of Business uses factor analysis in marketing research to construct perceptual maps, a graphic to display the perceptions of customers, and other product positioning devices. In the field of Physical Science, factor analysis is used for research in mineral analysis to identify factors that correspond to different mineral associations. Factor analysis is used in Education for construct validity, instrument development and more. These are just some of the types of research that are conducted in these fields that involve factor analysis.

There are essentially two types of Factor Analysis, Exploratory Factor Analysis and Confirmatory Factor Analysis. Exploratory Factor Analysis (EFA) is used when the researcher begins an analysis with no clear idea of what will be found. The researcher is actually exploring the data to find a structure that makes sense (Child, 1978). Exploratory factor analysis could often be referred to as theory generation rather than theory testing (Thorndike, 1997). The goal is really to identify the factors that underlie the data obtained in the research (Moser, 2004). On the other hand, Confirmatory Factor Analysis (CFA) is used when the researcher enters the analysis with some clear expectations on what will be found. Researchers are actually using the analysis to support their hypothesized theory (Thorndike, 1997). In confirmatory factor analysis, a more specific hypothesis about the factor structure is imposed by the researcher in the hopes that such a specific hypothesis will be supported by a given covariance structure. If the specific hypothesis is supported by the given data, then confirmatory factor analysis can provide self-validating information (Kim & Mueller, 1978). The researcher is explicitly building a model that states how the factors will contribute to the data obtained in the research (Moser, 2004).

Factor Analysis is a multivariate statistical technique that describes the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors. The factors are hypothetical constructs whose values can only be estimated from observed data (Johnson & Wichern, 1988). In scientific research, a construct is a type of concept used to describe events that share similar characteristics (Borg & Gall, 1989). It has been noted that the determination of the number of factors to retain is one of the most critically important decisions that a researcher makes in Factor Analysis (Zwick & Velicer, 1986). In Exploratory Factor Analysis (EFA), researchers do not know the true number of factors that are underlying the data. Therefore, their decision on the number of factors to retain can result in too few or too many. The decision that results in too few factors is considered an underestimation. The decision that results in too many factors is considered an overestimation. The decision to retain the correct number of factors, in an Exploratory Factor Analysis, is important because it is made prior to factor rotation. Factor rotation methods are utilized to find equivalent solutions that are easier to interpret (Moser, 2004). It has been noted that the original factor loadings may not be readily interpretable; therefore, it has become usual practice to rotate them until a simpler structure is achieved that is more interpretable (Johnson & Wichern, 1988). Consequently, the decision of retaining the correct number of factors can impact the factor rotation method, factor patterns, factor scores, and the interpretability of the factors. The interpretation of factors is based on the assumption that the correct number of factors is retained from a study (Turner, 1998).

Extraction of a different number of factors other than the true number of factors can dramatically affect the results of a study. In the case of underestimation, the loss of potentially important information is of the highest concern. It has been noted that underestimation of a data set results in a situation where the true number of factors in the data set cannot be accurately described (Cattell, 1978; Fava & Velicer, 1996; Velicer, Eaton, & Fava, 2000). When underestimation occurs, the variables that are not yet accounted for a true factor can be mistakenly represented by another factor. These types of errors are likely to lead a researcher to misinterpret the actual model that is defined from the data. Fabrigar, Wegener, MacCallum, and Strahan (1999) noted that variables that were not captured by their true factor may mistakenly appear too poorly defined by another factor. A study conducted by Joseph Fava and Wayne Velicer in 1996, showed that underestimation of the number of factors retained led to substantial degradation of scores in the estimation methods of Principal Component Analysis and Maximum Likelihood method.

Many researchers agree that overestimation of the factors is a less severe problem than the underestimation of factors. Fabrigar et al. (1999) and others agree that underestimation of factors can lead to more distorted outcomes than overestimation of factors. However, overestimation of factors is problematic and should be avoided in Exploratory Factor Analysis. Zwick and Velicer (1986) found that overestimation of the factors may include factors that are not easily interpretable. Also, they noted that the results may include factors that a researcher will not be able to replicate. This alone would have serious implications to a researcher trying to identify underlying constructs. The ability to gather more data for support of their initial research would be compromised by their overestimation of factors in the initial study. However, some research claims that overestimation of factors can be handled after rotation by discarding the trivial factors without changing the factors of substance (Fava & Velicer, 1992). Other research suggests that factor splitting may occur in the overestimation of factors when a researcher starts a rotational method or even the collapse of a factor (Gorsuch, 1983). The implications that arise from the collapse of a factor would be that a common factor that has importance to the study at hand could be missed.

Underestimation and overestimation of the factors in Exploratory Factor Analysis can have dire consequences on one’s research. Whether overestimation is preferable to underestimation is not the issue of discussion. The issue of discussion is which method of extraction will be reliable in identifying the true number of factors. That is why the extraction methods that are available to researchers need to be scrutinized for the subsequent tendencies that each method might display under certain conditions in a study. The conditions referred to range from sample size, to the number of variables, to the level of factor saturation. If the methods of extraction all worked as they were intended, we would not have such intense and extensive research involving the methods of extraction. Furthermore, the development of new methods of extraction would not be necessary.

The purpose of this study was to investigate which method of extraction in factor analysis retained the true number of factors. The study also investigated each method of extraction when there were outliers present in the data. This study will be informative to researchers who are conducting studies that deal with data reduction or detecting data structure. One of the main reasons that this study will be informative to researchers is due to the fact that real data is never exactly multivariate normal (Johnson & Wichern, 1988). However, regardless of the form of the researcher's parent population, the sampling distribution of multivariate statistics will be approximately normal due to the central limit effect (Johnson & Wichern, 1988). The question that faces many researchers is how many factors to retain in factor analysis. This is a crucial problem that confronts researchers even when the data set contains no outliers. There have been many studies that have ranked the methods for data sets that are multivariate normal with no outliers. Therefore, it is the intention of this study to rank the methods of extraction on data sets that do contain outliers of varying degrees. The research questions to be answered are:

1. Will there be a difference among the methods of factor extraction on data that contains no outliers?
2. Will there be a difference among the methods of factor extraction on data that does contain outliers?
3. Is there some joint usage of methods that prove to be the most logical and safe alternative when choosing the number of factors?
4. Will the degree of the outlier have varying effects on the methods of extraction?

These questions will be answered based on the data collected from a Monte Carlo study.

This Monte Carlo study utilizes correlation matrices generated from factor structures determined by the investigator. The population correlation matrices in this study was designed with three factors. The number of variables in the population correlation matrices was set at twenty and forty. The level of factor saturation was set to values ranging from 0.2 to 0.8. The sample sizes were selected as a function of the number of variables (V). The formulas used to derive the sample sizes were 2(V)+10, 5(V), and 7(V)+10. Therefore, the cases with twenty variables were analyzed with sample sizes of fifty, one hundred, and one hundred fifty. The cases with forty variables were analyzed with sample sizes of ninety, two hundred, and two hundred ninety.

A principal components analysis (PCA) and principal factor analysis (PFA) were performed on each of the three hundred sample correlation matrices for each scenario. This resulted in a total of eighteen hundred sample correlation matrices for all the scenarios. The number of factors to be retained was determined by each of the seven factor extraction methods: Bartlett's, K1, MAP76, Map00, PA, and LgmPA.

The methods of extraction was also tested on data containing an outlier. A random sample of fifty population correlation matrices were chosen from each scenario of three hundred population correlation matrices. To incorporate the outlier, an observation was chosen at random within each of the fifty population correlation matrices. After choosing the observation, the vector containing the observation was multiplied by a scalar in increments to move it further out. The squared Mahalanobis distance was utilized to discover the distance in which the methods of extraction began to fail.

Definitions

Data reduction is an analytical method that involves reducing the dimensionality of a data set by extracting a number of underlying factors that can account for the variability in the data set (StatSoft, 2008).

Common factor is a factor in which two or more variables are correlated and hence contribute to the observed correlations between these variables.

Communality is the proportion of variance that each item has in common with other items (StatSoft, 2008).

Construct is a type of concept used to describe events that share similar characteristics (Borg & Gall, 1989).

Construct validation a validation study in which the test user desires to draw an inference from the test scores to performances that can be grouped under the label of some particular psychological construct (Crocker & Algina, 1986).

Eigenvalues are the variance extracted by the factors**.**

Factors are hypothetical constructs whose values can only be estimated from observed data (Johnson & Wichern, 1988).

Factor analysis a statistical technique that describes the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors.

Linear dependent implies that one of the vectors can be written as a linear combination of the other vectors (Moser, 2004).

Mahalanobis distance provides an indication of whether or not an observation is an outlier with respect to the independent variable values due to the fact that it is the distance of a case from the centroid in the multidimensional space, defined by the correlated independent variables (StatSoft, 2008).

Maximum Likelihood Method is a data reduction technique that requires a probability model (PCA and PFA do not require such a model) to describe the data (Moser, 2004).

Monte Carlo study is a computer-intensive technique for assessing how a statistic will perform under repeated sampling. In Monte Carlo methods, the computer uses random number simulation techniques to mimic a statistical population (StatSoft, 2008).

Multivariate normal is an extension of univariate normal to fit vector observations.

Orthogonal implies that a ninety-degree angle exists between entities that are under discussion.

Outliers are atypical, infrequent observations; data points which do not appear to follow the characteristic distribution of the rest of the data (StatSoft, 2008).

Principal Component Analysis (PCA) has the general objectives of data reduction and interpretation achieved by explaining the variance-covariance structure through a few linear combinations of the original variables (Johnson & Wichern, 1988).

Principal Factor Analysis (PFA) is similar to principle component analysis except for the fact that it only uses the variability in an item that it has in common with the other items, PCA assumes that all variability in an item should be used.

Specific factor is a factor that does not account for correlations between variables and is uncorrelated with each common factor and the specific factor for different variables are uncorrelated with one another.

CHAPTER 2: LITERATURE REVIEW

History

According to Harmon (1976), factor analysis is regarded as having a starting point in 1904, when Charles Spearman's paper, "General Intelligence, Objectively Determined and Measured" was published in the *American Journal of Psychology*. Spearman was trying to define the construct "intelligence" by working with scores obtained in examinations. He noticed certain systematic effects in the matrix of correlations between scores in different subjects. Upon these results, he composed some of the basis for factor analysis through the well-known Two-Factor Theory (Child, 1975). A considerable amount of work ensued over the next twenty years on the psychological theories and mathematical foundations of factor analysis. Researchers had realized that Spearman's Two-factor Theory was not always adequate to describe a battery of psychological tests. Eventually, the concept of multiple factor analysis arose, which involved extracting several factors directly from the matrix of correlations among tests. L.L. Thurstone gave a presentation about multiple factor analysis that was a particularly thorough and systematic discussion of the rationale and computations of factor analysis (Thurstone, 1947). Although Thurstone is usually given the credit for multiple factor analysis, there are many others, such as J.C. Garnett, who contributed greatly to the work of multiple factor analysis. Thurstone openly admitted that the centroid method is a computational compromise for the principal factor method. Thurstone's most remarkable contribution was the generalization of Spearman's tetrad-difference criterion to the rank of the correlation matrix as the basis for determining the number of common factors (Harman, 1976). Because of the early association with constructs, such as intelligence, factor analysis was developed primarily by researchers interested in psychometric measurements.

Many arguments arose over the psychological interpretations of several early studies involving factor analysis, which was even confounded further with the lack of powerful computing capabilities. This lack of powerful computing capabilities impeded the initial development of factor analysis as a statistical method. But, with the onslaught of powerful personal computers, a renewed interest has ensued in the computational and the theoretical aspects of factor analysis. Many of the early techniques have been abandoned in the wake of recent developments. But, it must be noted that these recent developments have also resolved the early controversies that surrounded factor analysis. However, it is still true that each application must be examined on its own merits to determine its success (Johnson & Wichern, 1988).

Purpose

Factor analysis is used for a variety of purposes such as revealing patterns of interrelationships among variables, detecting clusters of variables, and reducing a large number of variables to a smaller number of statistically independent variables that are each linearly related to the original variables. Essentially, the purpose of factor analysis is to describe the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors. The factors are hypothetical constructs whose values can only be estimated from observed data (Johnson & Wichern, 1988). In scientific research, a construct is a type of concept used to describe events that share similar characteristics (Borg & Gall, 1989). Factor analysis is one of several methods that is used in education and the social sciences for construct validation (Crocker & Algina, 1986). However, in factor analysis, it is believed that each construct is responsible for the observed correlations. The factors, in general, are merely convenient descriptive summarizations of the observed data. One of the main present day controversies is whether or not the factors have any real existence and causal, rather than just statistical implications. Nonetheless, the question of existence does not have to be established before a model can be used (Harman, 1976). Perhaps, an easier outlook on factor analysis can be summed up as having the aim to summarize the interrelationships among the variables in a concise but accurate manner as an aid in conceptualization (Goruch, 1974). Factor analysis has been considered an extension of principal component analysis because both methods attempt to approximate the covariance matrix, but the factor analysis model is more elaborate (Johnson & Wichern, 1988).

Factor Analysis in Education

The importance of factor analysis in the field of education can be traced to its creator, Charles Spearman, and his use of it to define and measure the construct of intelligence for his theory of intelligence (Johnson & Wichern, 1988). Of course, the early developments of factor analysis were nurtured by those interested in psychometric measurement. However, as previously stated, the importance of this type of analysis cannot be denied a place in the field of education. The construct of intelligence itself has a large role in the field of education if one is to measure the ability of the subject at hand. Robert Thorndike noted that factor analysis has grown to be one of the most used data-analytic procedures used in education and psychology. He further went on to mention that readers of school or clinical psychology literature are certain to encounter articles and studies that involve factor analysis (Thordike, 1997).

Factor analysis involves some very complicated mathematics (Thorndike, 1997) but its use has risen in the field of education due to our modern world of computers. Test validation is one area in education where factor analysis and correlation methods are the essential statistical techniques used (Crocker & Algina, 1986). The central issues in test validation are that both score meaning and the value implications of the scores are a basis for action (Messick, 1990). For instance, suppose a researcher is administering an achievement test to a large group of subjects that measures thirty different variables. Factor analysis can be used to determine whether each variable measures an individual type of achievement or whether two or more variables measure the same type of achievement (Borg & Gall, 1989). This knowledge will allow the researchers to focus on groups of variables that contribute to the measurement of the same type of achievement, which will inevitably ensure that their assessment of that type of achievement will be valid.

Factor analysis was also used in a study to develop an instrument to measure school climate at the secondary level of education. The instrument, called the Rutgers Organizational Climate Description Questionnaire for Secondary Schools (OCDQ-RS), was developed by Robert B. Kottkamp and others. In their study, a large pool of items designed to measure various aspects of school climate was reduced to five subscales that measure five dimensions of school climate (Borg & Gall, 1989; Kottkamp, Mulhern, & Hoy, 1987). Factor analysis was a key statistical technique that enabled the researchers to reduce the large number variables to just five factors, which they referred to as a subscale. However, the study was destined to use factor analysis again. Kottkamp and the other researchers were interested in seeing if the five subscales could be grouped into a smaller set of factors. On the second application of factor analysis, they were able to represent the five variables representing the five subscales by a smaller number of factors. This smaller number of factors was actually two factors in which they labeled as openness and intimacy (Borg & Gall, 1989; Kottkamp et al., 1987).

Through the use of factor analysis the researchers were able to reduce a large number of variables down to a few factors by combining the variables that are correlated with one another. The two factors from the Kottkamp study can be treated as variables in which each student could be given a factor score on each factor. Subsequent statistical analyses could be carried out on the factor scores from each student in order to answer various questions relevant to a study on school climate. For instance, a simple t-test could be performed to find the difference in perception from parochial and public schools on factor one (Borg & Gall, 1989).

In 2009, a study was conducted in the Netherlands on the relationship between existential fulfillment and burnout among secondary school teachers. A confirmatory factor analysis was utilized in this study which revealed a three-dimensional construct with interdependent dimensions. The study confirmed the hypothesis concerning negative relationships between the existential fulfillment dimensions on the one hand and the burnout dimensions exhaustion and cynicism on the other. It also confirmed a positive relationship between existential fulfillment dimensions and the burnout dimension professional efficacy. The study established the importance of existential fulfillment for the prevalence and prevention of burnout among secondary teachers (Loonstra, Brouwers, & Tomic, 2009).

In Canada, a study was conducted on the relationship between school engagement and dropouts. The concept between the two figures prominently in school dropout theories, but little empirical research has been conducted on its nature and course. The importance of this research would benefit those individuals interested in preventing school alienation during adolescence, which in turn might reduce the growing number of dropouts in future generations. Through the use of factor analysis and structural equations the researchers were able to use global engagement reliability to predict school dropout (Archambault, Janosz, Fallu, & Pagani, 2009).

A study conducted by Fred N. Kerlinger and Elazar J. Pedhazer utilized factor analysis to help explain how attitudes toward education and perceptions of desirable traits of teachers are related. The study was conducted over several states using results from over three thousand teachers and graduate students of education. The raw data from the subjects came from three education attitude scales and four teacher trait perception instruments. In the study, the researchers used second order factor analysis on the correlations among the factors obtained by the first order factor analysis. The researchers felt that the items of the educational attitude and trait perception obtained from the first order factor analysis would yield two second order factors. Their hypothesis was confirmed when the second order factor analysis showed that progressive attitudes and person-oriented teachers perceptions fell on one factor, and the second factor was composed of traditional attitudes and task-oriented perceptions. The researchers established the basic hypothesis that individuals with progressive attitudes towards education, in selecting traits they feel as desirable in teachers, choose person-oriented traits, traits that are in accordance with progressive education beliefs. Also, well established was the fact that individuals with traditional attitudes towards education choose task-oriented traits, traits that are in accordance with traditional educational beliefs (Kerlinger & Elazar, 1968).

From the previous studies, we can now fully understand that factor analysis is a very valuable tool in educational research. However, with all statistics one must be very careful in its application. An old adage used in the field of statistics states “Garbage in, garbage out,” (Borg & Gall, 1989, p. 623). Applying this old adage to factor analysis simply means that the factors extracted are only as interpretable as the variables entered. In laymen’s terms this means that if the variables have little or no conceptuality in common, then interpretation of the factors extracted will have little to no meaning. Therefore, the number and type of variables entered into a factor analysis should be carefully considered by the researcher to ensure that the results will have legitimate meaning (Borg & Gall, 1989; Johnson & Wichern, 1988).

Orthogonal Factor Model

In factor analysis, we start out with a data set of p variables and n observed values. The data set can be denoted by a n x p matrix, such as:

$$X = \begin{matrix} X_{11} & X_{12} & X_{13} \ldots\ldots\ldots\ldots X_{1p} \\ X_{21} & X_{22} & X_{23} \ldots\ldots\ldots\ldots X_{2p} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ X_{n1} & X_{n2} & X_{n3} \ldots\ldots\ldots\ldots X_{np} \end{matrix}$$

The objective of factor analysis is to represent a variable x, corresponding to the columns of X, in terms of several underlying factors. The basic factor analysis model is:

$$X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + \ldots\ldots\ldots\ldots + l_{pm}F_m + \varepsilon_p$$

Or, in matrix notation,

$$X = L\,F + E$$

The factor model postulates that X is linear dependent on some unobservable random variables F_1, F_2, …… F_m, and p additional sources of variation ε_1, ε_2, …… ε_p. The F variables are referred to as common factors and the ε’s are referred to as specific factors

(Johnson & Wichern, 1988). The common factor is a factor in which two or more variables are correlated and hence contribute to the observed correlations between these variables. The specific factor is uncorrelated with each common factor and the specific factor for different variables are uncorrelated with one another. Therefore, specific factors do not account for correlations between variables (Crocker & Algina, 1986). The coefficient l is frequently referred to as the "loadings." Without any loss of generality, it is assumed that the F's and the ε's have zero means and unit variances. This is due to the fact that they are unknown in practice. The n unique factors are supposed to be independent of one another and also independent of the m common factors (Johnson & Wichern, 1988; Rencher, 1995).

EFA vs. CFA

Exploratory Factor Analysis (EFA) is used when the researcher begins an analysis with no clear idea of what will be found. The researcher is actually exploring the data to find a structure that makes sense (Child, 1978). Exploratory factor analysis could often be referred to as theory generation rather than theory testing (Thorndike, 1997). Jae-On Kim and Charles W. Mueller (1978) described four steps in applying exploratory factor analysis to actual data in the series: Quantitative Applications in the Social Sciences. The four basic steps they presented are:

1. The data collection and preparation of the relevant covariance matrix.
2. The extraction of the initial factors.
3. The rotation to a terminal solution and interpretation.
4. Construction of factor scales and their use in further analysis.

Confirmatory Factor Analysis (CFA) is used when the researcher enters the analysis with some clear expectations of what will be found. The researcher is actually using the analysis to support their hypothesized theory (Thorndike, 1997). In confirmatory factor analysis, a more specific hypothesis about the factor structure is imposed by the researcher in the hopes that such a specific hypothesis will be supported by a given covariance structure. If the specific hypothesis is supported by the given data, then confirmatory factor analysis can provide self-validating information (Kim & Mueller, 1978). The basic steps for conducting a confirmatory factor analysis would be similar to Kim and Mueller's approach to EFA with the steps preceded by:

1. Model specification.

2. Determination of model identification.

Furthermore, the steps from Kim and Mueller should be followed by assessment of the models fit. If an unacceptable model fit is found in confirmatory factor analysis, exploratory factor analysis can be preformed to find underlying constructs for the set of measured variables (Child, 1978).

There are similarities that exist between exploratory and confirmatory factor analysis other than the fact that both are powerful multivariate statistical techniques. For example, both statistical techniques are based on linear statistic models. Furthermore, both statistical procedures are used to identify latent constructs that might be represented by a set of measured variables (Moser, 2004).

The differences between the two statistical procedures arise in the application process that the two procedures are to use. In exploratory factor analysis, the researcher decides the number of factors by examining the output of the data. From this point the researcher determines the factor structure (model). Exploratory factor analysis allows all items to load on all factors. In confirmatory factor analysis, the researcher specifies the model and the number of factors prior to running the analysis. The researcher is actually specifying which items load on which factors. The researcher is hoping that the factor structure fits the model in which he/she has specified (Kim & Mueller, 1978).

The study conducted in this paper will utilize exploratory factor analysis, because in exploratory factor analysis data are simply explored and information provided on how many factors are required to represent the data. This study is concentrating on factor extraction, the number of factors to retain. No model specification or control of the factor loadings will be necessary to explore the techniques of factor extraction. Nor will there be a need to explore the techniques of factor extraction with data containing outliers of varying degrees. Of course, the study's simulation will provide random data sets with a specified number of factors, but through this randomization one will not have a preconceived notion of which items will load on which factors.

Estimation Methods

There are many methods for parameter estimation in factor analysis. The three most commonly used are Principal Component Analysis, Principal Factor Analysis, and the Maximum Likelihood Method (Johnson & Wichern, 1988). The goal of Principal Component Analysis (PCA) is to summarize patterns of correlations among observed variables and to reduce a large number of observed variables to a smaller number of factors. This is accomplished by seeking a linear combination of variables in such a way that the maximum variance is extracted from the variables. This first linear combination is called the first component and the variance of this first component is equal to the largest eigenvalue in the covariance matrix. The next step in PCA is to remove this variance and search for a second linear combination, uncorrelated with the first component, which explains the maximum proportion of the remaining variance. This would be called the second component and the variance of this second component is equal to the second largest eigenvalue in the covariance matrix. These steps are repeated until all variances are accounted for in the set of variables. PCA transforms a set of correlated variables into a set of uncorrelated components. The intention of Principal Component Analysis is to have a smaller number of components that account for most of the variance of the original set of variables (Stevens, 1986).

Principal Factor Analysis (PFA) is a modification of the Principal Component Analysis (Johnson & Wichern, 1988). It is sometimes referred to as principal axis factoring. PFA seeks the least number of factors, which can account for the common variance of a set of variables. PFA differs from Principal Component Analysis in that estimates of communality are in the positive diagonal of the observed correlation matrix. These estimates of communality are derived through an iterative procedure with the squared multiple correlations of each variable with all other variables used as a starting point in the iterations (Tabachnick & Fidell, 2001). The goal of PFA is to extract maximum orthogonal variance from the data set with each succeeding factor.

Principal Component Analysis and Principal Factor Analysis will lead to similar conclusions for most data sets (Wilkinson, Blank, & Gruber, 1996). However, Snook and Gorsuch (1989) conducted a Monte Carlo study to determine whether PCA or PFA was more accurate and at what point did the results merge to have equivalent findings. They found that PFA was better suited to data sets that had a low number of variables than PCA. Also, in the study they indicated that the two methods should not be considered to have equivalent finding until forty or more variables are present in the matrix (Snook &

Gorsuch, 1989). A Monte Carlo study uses computer generated data sets to mimic statistical populations according to the researcher's prescription. PCA is generally preferred for purposes of data reduction, while PFA is generally preferred when the research purpose is detecting data structure. In 1998, the following statement on the pros and cons of PCA and PFA were given by G. David Garson:

- PCA determines the factors which can account for the total (unique and common) variance in a set of variables. This is appropriate for creating a typology of variables or reducing attribute space. PCA is appropriate for most social science research purposes and is the most often used form of factor analysis.

- PFA determines the least number of factors which can account for the common variance in a set of variables. This is appropriate for determining the dimensionality of a set of variables such as a set of items in a scale, specifically to test whether one factor can account for the bulk of the common variance in the set, though PCA can also be used to test dimensionality. PFA has the disadvantage that it can generate negative eigenvalues, which are meaningless (www2.chass.ncsu.edu/garson/pa765/factor.htm).

Maximum likelihood is another method based on linear combinations of the variables to form factors. The maximum likelihood approach requires a probability model to describe the data (Moser, 2004). Also, multivariate normality is essential for maximum likelihood to estimate the factor loadings and specific variances (Johnson & Wichern, 1988). An advantage of maximum likelihood is that it generates a chi-square goodness-of-fit test that allows the researcher to increase the number of factors one at a time until a satisfactory goodness-of-fit is obtained. But, a disadvantage of the chi-square goodness-of-fit test is that it can lead to an overestimation of factors due to its sensitivity in large samples (Garson, 1998).

Factor Extraction

A principal difficulty that arises in factor analysis is one that relates to the choice of the number of factors. This is a critical point in the research, in which the researcher needs to carefully consider the data and use his/her best judgment. It is very important for the researcher to remember the advantages and the limitations of the various decision rules and make a well-reasoned decision based on the nature of the analysis (Hetzel, 1995). Making the incorrect choice may lead to under- extraction of the factors, which usually equates to a loss of information. Overestimation of the factors will tend to lead researchers to include random variation in the data, which will have affect interpretation later in the study (Zwick & Velicer, 1986). It has been noted by some researchers that underestimation can lead to more distorted results than overestimation (Fabrigar et al., 1999; Tinsley & Tinsley, 1987). Researchers have developed a number of ways to extract the correct number of factors. Some of the most commonly considered tests for determining the number of factors are the Guttman-Kaiser rule, Scree test, Bartlett's test, Minimum Average Partial, and Parallel Analysis.

The Guttman-Kaiser rule is a commonly used method that uses all factors that have an eigenvalue greater than one. The eigenvalue is the sum of the squared loading values for a factor that shows the amount of variance a factor can account for. It has been shown in several studies that the Guttman-Kaiser rule has a strong tendency to overestimate the number of factors. For instance, R.L. Linn performed a Monte Carlo study of the Guttman-Kaiser rule based on seven predetermined factors, twenty and forty variables, and sample sizes of one hundred and five hundred. He found that the underestimation was minor, but the overestimation occurred approximately 66% of the time (Linn, 1968). Another study, conducted by Zwick and Velicer, showed that the Guttman-Kaiser rule overestimated the correct number of factors most of the time (Zwick & Velicer, 1982; 1986). The majority of the studies showed overestimation of the factors, but there has been the report of underestimation of factors when utilizing the Guttman-Kaiser rule (Tinsley & Tinsley, 1987).

The Scree test, developed by R.B. Cattell, is another test in determining the number of factors. The Scree test is a graphical method for determining the number of factors. This is accomplished by plotting the eigenvalues in the sequence of the principal factors. The number of factors to be retained are chosen by their position on the graph. All factors that lie above the point where the plot levels off to a linear decreasing pattern are retained (Cattell, 1978). The Scree test has been shown in several studies to be inaccurate in determining the correct number of factors. Zwick and Velicer found that it was accurate about half of the time with a large tendency to overestimate (Zwick & Velicer, 1982; 1986). Another study by Linn, Tucker, and Koopman in 1969 found the test to be about 67% accurate (Tucker, Koopman, & Linn, 1969). It was noted in one study that the determination of where the plot levels off to a linear decreasing pattern can have an uncertainty depending on the actual graph itself (Tanguma, 1999). Tanguma noted that the graph could have a gradual slope with no obvious linear break, or have more than one point to construct a linear break.

M. Bartlett developed a statistical test of the null hypothesis, following D. Lawley's test for maximum likelihood factor analysis, that the remaining p-m eigenvalues are equal. The test sequentially excludes each eigenvalue until the approximate chi-square test of the null hypothesis of equality fails to be rejected. The first m components are retained in this test (Bartlett, 1950; 1951). It was noted in several studies that it tended to overestimate the number of factors (Zwick & Velicer, 1982; 1986). This should be expected due to the fact that the maximum likelihood test shows a consistent tendency to overestimate the true number of factors (Glorfed, 1995).

Parallel analysis, developed by John L. Horn, is another test for determining the number of factors. It is a sample-based adaptation of the population based Guttman-Kaiser rule (Horn, 1965). Horn noted that sample correlation matrices that were generated from the population identity matrix had off-diagonal elements that assumed random correlations larger in absolute value than zero. These correlations resulted in matrices with initial eigenvalues greater than one, whereas the final eigenvalues were less than one. Horn stated that a number of correlation matrices of p uncorrelated normal random variables with a sample size n, where n and p are the same as the corresponding entries in the data set under study, be constructed and their eigenvalues averaged. These averaged eigenvalues would be compared to the eigenvalues from the real data correlation matrix. The only factors that would be considered in further analysis would be factors corresponding to actual eigenvalues that exceed the average eigenvalues. He noted that actual eigenvalues equal to or less than the averaged random eigenvalues would be considered as due to random sampling variability (Glorfed, 1995). Numerous studies have shown that Horn's Parallel Analysis is very accurate in determining the number of factors. Zwick and Velicer reported in their study that the Parallel Analysis determined the correct number of factors a very large percentage of the time (Zwick & Velicer, 1982; 1986). If the correct number of factors was not determined, it was noted that Parallel Analysis tended to overestimate the factors a majority of the time (Zwick & Velicer, 1982; 1986). A study conducted on ecological data found Parallel Analysis as an efficient and robust means for determining the number of components (factors) when used in conjunction with Principal Components Analysis (Franklin, Gibson, Robertson, Pohlmann, & Fralish, 1995).

Longman's Parallel Analysis, developed by Stewart Longman and others, is a regression equation for predicting parallel analysis values used to decide the number of factors to retain in factor analysis. Longman and others believed that Horn's Parallel Analysis was too dependent on chance. Therefore, their regression equation uses the 95^{th} percentile point in the distribution of eigenvalues generated from random data matrices (Longman, Cota, Holden, & Fekken, 1989). Eigenvalues from ones research that is greater than the 95^{th} percentile eigenvalues generated from Longman's regression equations are retained. According to Skinner (1989), the 95^{th} percentile eigenvalues represent a benchmark for identifying factors that may have been extracted by mere chance.

The Minimum Average Partial, developed by Wayne Velicer, is a method that is based on the matrix of partial correlation. After each factor has been taken out, the average of the squared partial correlation is calculated. No further factors are extracted when the minimum average squared partial correlation is obtained. The minimum average partial correlation is obtained when the residual matrix closely resembles an identity matrix (Zwick & Velicer, 1982; 1986). In 2000, the Minimum Average Partial was revised with the partial correlations raised to the fourth power instead of the second power (Velicer et al., 2000).

CHAPTER 3: METHODS

Method

It is the intention of this study to rank the methods of extraction on data sets that do not contain outliers and on data sets that contain outliers of varying degrees. The research questions to be answered are:

1. Will there be a difference among the methods of factor extraction on data that contains no outliers?
2. Will there be a difference among the methods of factor extraction on data that does contain outliers?
3. Is there some joint usage of methods that prove to be the most logical and safe alternative when choosing the number of factors?
4. Will the degree of the outlier have varying effects on the methods of extraction?

To answer these questions, this study investigated the performance of the seven factor extraction procedures using Monte Carlo methods. The Monte Carlo methods generated random samples of data under known and controlled population conditions. The population correlation matrices that were randomly generated varied with reference to the particular aspects of interest involved in the study. The factor loading patterns underlying these random population correlation matrices were constructed to reveal clear simple structure for ease of assessment. The number of common factors, the number of variables, the number of observations and the level of communality were controlled for these randomly generated population correlation matrices. It should be noted that more than one population correlation matrix can be generated having the desired number of variables, factors, and level of communality. Each of the randomly generated population correlation matrices were then analyzed using the maximum likelihood method. In related research, the maximum likelihood method was employed to determine if the assumed number of factors was correct (Zwick and Velicer, 1982).

Generation of a Population Correlation Matrix

In the Monte Carlo study, a population correlation matrix was generated under the assumption that the common factor will hold true in the population. Each population correlation matrix was determined under the following conditions using SAS/IML version 9.1.3. First, a population matrix (L) was created in accordance with the number of variables, the number of factors, and the level of factor saturation under consideration in the study.

```
L=(J(nrow,ncol,value)//J(nrow,ncol,value)//J(nrow,ncol,value)||
   (J(nrow,ncol,value)//J(nrow,ncol,value)//J(nrow,ncol,value)||
   (J(nrow,ncol,value)//J(nrow,ncol,value)//J(nrow,ncol,value);
```

This matrix can also be referred to as the matrix defining the factors. Second, population matrix (L) was multiplied by its transpose (L`) to give the resulting covariance matrix (R).

```
LLPrime=L*L`;
R=LLPrime;
```

Third, a substitution of ones was employed into the diagonal of the covariance matrix (R) to produce a correlation matrix raised to its full rank.

```
P=ncol(R);
Do I=1 to P;
R[I,I]=1;
```

Fourth, generation of multivariate normal data was accomplished by multiplying Cholesky’s decomposition, ROOT(), by the random normal variable generator RANNOR. The RANNOR generates two variables that are independent random samples from a normal distribution having a mean of zero and a standard deviation of one (Fan, Felsovalyi, Sivo, & Keenan, 2001). The J function, in this part of the program, controls the number of observations by placing the desired value in place of nrow.

```
Z=Root (R);
X=J(nrow,ncol,value);
Y=Rannor(X)*Z;
```

It should be noted that this Monte Carlo method of generating population correlation matrices allowed the researcher to control the number of factors, the number of variables, the number of observations, and the level of factor saturation.

Design of the Monte Carlo Study

This Monte Carlo study utilized correlation matrices generated from factor structures determined by the investigator. This gave the advantage of known criterion in which to judge each test for its accuracy. The population correlation matrices in this study were designed with three factors. The number of variables in the population correlation matrices were set at twenty and forty. In present day work, this should be considered a relatively small to moderate data set, since many real data sets can have as many as two hundred variables. The level of factor saturation was set to values ranging from 0.2 to 0.8. The range is consistent with levels that have been used in previous simulation studies. The sample sizes were selected as a function of the number of variables (V). The formulas used to derive the sample sizes are 2(V)+10, 5(V), and 7(V)+10. The sample sizes were selected in this manner in order to have a resemblance to applied usage. Therefore, the cases with twenty variables will be analyzed with sample sizes of fifty, one hundred, and one hundred fifty. The cases with forty variables will be analyzed with sample sizes of ninety, two hundred, and two hundred ninety. It has been noted that sample sizes of this nature appear to include a representative range of sample sizes reported in applied educational and psychological research (Zwick and Velicer, 1984).

A principal components analysis (PCA) and principal factor analysis (PFA) was performed on each of the three hundred sample correlation matrices for each scenario. This resulted in a total of eighteen hundred sample correlation matrices for all the scenarios. The number of factors to be retained was determined by each of the seven factor extraction methods: BART, K1, MAP76, Map00, PA, and LgmPA.

Generation of an Outlier

To test the methods of factor extraction on data containing outliers, fifty randomly chosen population correlation matrices were selected from each of the scenarios containing the original three hundred population correlation matrices. Therefore, a total of three hundred population correlation matrices were selected from the grand total of eighteen hundred population correlation matrices. For each of the randomly selected population correlation matrices, a corresponding correlation matrix was generated from the originally derived correlation matrix with an outlier incorporated into the data. The theory behind this process was to incorporate the outlier to a randomly chosen observation from the second generated correlation matrix with a different mean. To achieve this difference in means, a scalar was multiplied to each randomly selected observation to ensure that the mean was considerably different from the mean of the original observation. The multiplication of the scalar was incorporated in a SAS macro program using proc IML. This program allowed the researcher to actually create five correlation matrices at a time with the outlier of varying degrees incorporated into the original correlation matrix. All five population correlation matrices were saved under a different name so further analysis could be implemented. The only structure that was changed in the five correlation matrices was the randomly selected observation. Thus, an outlier was formed into the original randomly selected correlation matrix. The multiplication of the scalar was sometimes repeated as many as twenty times in order to achieve the desired effects. The variance-covariance structure will stay the same, only the mean will be changed. For example, a correlation matrix with twenty variables with a sample size of fifty will have a composition of ninety-eight percent of the observations from one normal and two percent will come from another normal. The outlier will change the correlation structure of the matrix, but the question to be answered is whether the methods of extraction will still identify the same number of factors. A principal components analysis and a principal factor analysis were performed on each of the fifty population correlation matrices with the outlier to see which method of factor extraction is affected by the outlier.

To identify the point where the methods of extraction tend to fail, the Mahalanobis distance was calculated on each outlier. The Mahalanobis distance is a statistical distance standardized by standard deviation along principal components (Moser, 2004). The Mahalanobis distance of a multivariate vector $x=[x_1, x_2, x_3,\ldots, x_n]'$ and $\mu =[\mu_1, \mu_2, \mu_3,\ldots, \mu_n]'$ having a covariance matrix $\sum$, can formally be illustrated as (Moser, 2004; Johnson & Wichern, 1988):

$$D(x,\mu) = [(x - \mu)'\textstyle\sum^{-1}(x - \mu)]^{1/2}$$

or

$$D^2(x,\mu) = (x - \mu)'\textstyle\sum^{-1}(x - \mu)$$

The Mahalanobis distance was computed from each observation to the mean. This was accomplished through SAS programming using PROC PRINCOMP with the STD option. This produced principal component scores having an identity covariance matrix in the resulting data set. At this point Euclidean distance and Mahalanobis distance are equal until a data step was inserted defining the Mahalanobis distance to complete the required distance (SAS Institute Inc., 2004). The SAS program used in calculating the Mahalanobis distance was coded to find the Mahalanobis distance (D) and the Mahalanobis distance squared (D^2).

The Scree test, plots of the eigenvalues, were analyzed for every analysis performed. The plots were examined by a rater that was briefed on the definition of a Scree plot and shown several known examples of actual Scree plots. The definition and explanation of the Scree test was taken from a book, *A Step-by-Step Approach to Using SAS® for Factor Analysis and Structural Equation Modeling*, by Larry Hatcher. The rater was a college graduate who holds an undergraduate degree in mathematics and a master's degree in the field of experimental statistics. The graphs themselves were cut and pasted from the SAS output onto 8½" by 12" sheets of paper. The same plots were compared to the researcher's own decision on the number of factors to retain. No significant difference was found between the rater and the researcher at an alpha level at 0.05.

Results

The mean number of factors retained by each method of factor extraction in each scenario was computed. The mean for each was then subtracted from the population criterion of three. A positive difference in scores indicates an overestimation of the population value of three, while a negative difference indicates an underestimation. Table 1 presents a summary of the results for principal components analysis (PCA) and principle factor analysis (PFA) side by side according to the number of variables, twenty, and sample size, fifty. A detailed description of Table 1 will be given. However, due to the fact that Tables 2-6 follow the exact same format, a detailed description will be omitted.

The first column in Table 1 inspects the performance of the seven methods of factor extraction for PCA and PFA. Under PCA, Bart had a mean average of 3.08333 factors and, thus had a mean difference of 0.08333, an overestimation. K1 had a mean average of 3.0 and a mean difference of 0.0 under both estimation methods of PCA and PFA. In PCA, LgmPa had a mean difference of 2.76 with a mean difference of -0.24, an underestimation. LgmPa also showed an underestimation in PFA with a mean of 2.6 with a mean difference of -0.4. The Map00 had a mean of 3.053333 with a mean difference of 0.053333, an overestimation in PCA. In PFA, the Map00 showed a slightly greater overestimation than it displayed in PCA with a mean of 3.05667 and a mean difference of 0.5667. The MAP76 showed similar results between the two estimation methods as did the MAP00 with overestimation slightly increasing from .03333 in PCA to 0.05 in PFA. PA had a mean of 2.88 with a difference of -0.12, an underestimation in PCA. However, the factor extraction PA only slightly underestimated in PFA with a mean of 2.97 and mean difference of -0.03. The SCREE overestimated the most out of all seven methods in both PCA and PFA with a corresponding difference of 0.17667 and 0.27333. In PFA, the SCREE actually overestimated more than any other extraction method for both estimation methods.

Table 1 - Mean Number of Factors Retained and Mean Difference. (V = 20 N = 50)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.08333	0.08333		
K1	3.0	0.0	3.0	0.0
LgmPA	2.76	-0.24	2.6	-0.4
MAP00	3.05333	0.05333	3.05667	0.05667
MAP76	3.03333	0.03333	3.05	0.05
PA	2.88	-0.12	2.97	-0.03
SCREE	3.17667	0.17667	3.27333	0.27333

Table 2 was composed with the same form as Table 1. The difference in the tables is distinguishable by the increase to one hundred in the sample size. Under these conditions BART and MAP00 tended to slightly overestimate while the SCREE tended to moderately overestimate with the PCA estimation method. LgmPA and PA tended to slightly underestimate while K1 and MAP76 performed perfectly. Using the PFA method of estimation, the factor extraction methods K1, MAP00, MAP 76, and PA preformed perfectly. LgmPA moderately underestimated, while the SCREE moderately overestimated. It seems as if the larger sample size worked well in conjunction with the PFA method of estimation and several methods of factor extraction.

Table 2 - Mean Number of Factors Retained and Mean Difference. (V = 20 N = 100)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.07333	0.07333		
K1	3.0	0.0	3.0	0.0
LgmPA	2.95	-0.05	2.74	-0.26
MAP00	3.00333	0.00333	3.0	0.0
MAP76	3.0	0.0	3.0	0.0
PA	2.99333	-0.00667	3.0	0.0
SCREE	3.15	0.15	3.11	0.11

Table 3 is also presented in the same manner as Tables 1 and 2. The sample size in this scenario was one hundred fifty. The PCA estimation method had K1, MAP00, MAP76 and PA performing perfectly. LgmPA slightly underestimated while the SCREE grossly overestimated. It would appear that the increase in sample size also positively impacted the precision of several of the extraction methods in PCA. The PFA estimation method had the same results as the previous scenario with a sample size of one hundred. The only change that resulted was that the SCREE grossly overestimated even more. However, the LgmPA did show a little improvement, but still showed moderate underestimation.

Table 3 - Mean Number of Factors Retained and Mean Difference. (V = 20 N = 150)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.05	0.05		
K1	3.0	0.0	3.0	0.0
LgmPA	2.97666	-0.02334	2.75333	-0.24667
MAP00	3.0	0.0	3.0	0.0
MAP76	3.0	0.0	3.0	0.0
PA	3.0	0.0	3.0	0.0
SCREE	3.35667	0.35667	3.47667	0.47667

The results in Table 4, 5, and 6 closely parallel the results in Tables 1, 2, and 3. The most notable change between the results is that of LgmPA and PA. Both of those methods of extraction performed perfectly in all three scenarios under both estimation methods. It would seem to indicate that the increase in variables and sample size had a positive impact on how those methods performed.

Table 4 - Mean Number of Factors Retained and Mean Difference. (V = 40 N = 90)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.12333	0.12333		
K1	3.0	0.0	3.0	0.0
LgmPA	3.0	0.0	3.0	0.0
MAP00	3.11	0.11	3.11	0.11
MAP76	3.02667	0.02667	3.02667	0.02667
PA	3.0	0.0	3.0	0.0
SCREE	3.14	0.14	3. 06667	0.06667

Table 5 - Mean Number of Factors Retained and Mean Difference. (V = 40 N = 200)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.04	0.04		
K1	3.0	0.0	3.0	0.0
LgmPA	3.0	0.0	3.0	0.0
MAP00	3.01333	0.01333	3.00333	0.00333
MAP76	3.0	0.0	3.0	0.0
PA	3.0	0.0	3.0	0.0
SCREE	3.02	0.02	3. 12	0.12

Table 6 - Mean Number of Factors Retained and Mean Difference. (V = 40 N = 290)

	PCA		PFA	
	Mean	d	Mean	d
BART	3.02667	0.02667		
K1	3.00333	0.00333	3.0033	0.0033
LgmPA	3.0	0.0	3.0	0.0
MAP00	3.00333	0.00333	3.00333	0.00333
MAP76	3.0	0.0	3.0	0.0
PA	3.0	0.0	3.0	0.0
SCREE	3.00667	0.00667	3. 06	0.06

Table 7 presents a one-way analysis of variance procedure to test whether the means are equal for all seven factor extraction methods (H_0: $\mu_{Bart} = \mu_{K1} = \mu_{LgmPA} = \mu_{MAP00} = \mu_{MAP76} = \mu_{PA} = \mu_{SCREE}$) or whether there exist some difference among the methods (H_a: At least one inequality). Since the null hypothesis was rejected in Table 7, we can conclude that there is at least one inequality. To decide which methods of extraction are different from one another, a Tukey's multiple comparison procedure is included. Means with the same letter designated by the Tukey procedure indicates which extraction methods are not significantly different. Those extraction methods with different letters designated by the Tukey procedure indicate a significant difference. For instance, the ANOVA in Table 7 indicates that the null hypothesis is rejected with a p-value < 0.0001. Therefore, we can conclude that there is at least one significant inequality among the means. Tukey's multiple comparison procedure shows that the SCREE, PA, and LgmPA are different from each other and all other methods of extraction using PCA under the conditions involving twenty variables with a sample size of fifty. Tukey's procedure also shows that BART, MAP00, and MAP76 are not significantly different for this scenario. Furthermore, it shows that MAP00, MAP76, and K1 are not significantly different.

Table 7 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 20 N = 50)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	34.2257143	5.7042857	56.92	<.0001
Error	2093	209.7666667	0.1002230		
Corrected Total	2099	243.9923810			

R-Square	Coeff Var	Root MSE	NUM Mean
0.140274	10.55937	0.316580	2.998095

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	34.22571429	5.70428571	56.92	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	TEST
	A	3.17667	300	SCREE
	B	3.08333	300	BART
	B			
C	B	3.05333	300	MAP00
C	B			
C	B	3.03333	300	MAP76
C				
C		3.00000	300	K1
	D	2.88000	300	PA
	E	2.76000	300	LgmPA

Tables 8-18 give the same information as Tables 7 for each corresponding scenario. If the null hypothesis was rejected in the given scenario, then the ANOVA is followed by Tukey's multiple comparison procedure to show which methods of extraction are significantly different and which are not significantly different. The only difference in Tables 7-18 is the null hypothesis under the PFA estimation methods do not have μ_{BART} included. This is due to the fact that Bartlett's chi square can only be run through principal components analysis. However, the other six methods of extraction were tested.

Table 8 presents the PFA method of estimation for twenty variables and a sample size of fifty. The one-way analysis of variance procedure indicates that the null hypothesis was rejected. Since the null hypothesis was rejected, we can conclude that there is at least one inequality. Tukey's multiple comparison procedure shows that the MAP00, MAP76, and K1 are not significantly different. However, there was a slight overestimation of the factors with MAP00 and MAP76. Tukey's also shows that there is no significant difference between K1 and PA. On the other hand, the PA method of extraction shows a slight underestimation of the factors. These test are not significantly different, the fact that some tend to overestimate while other tend to underestimate the true number of factors is important to the researcher so he/she will know the tendencies of that particular method of extraction in this type of scenario. The Scree displayed a moderate overestimation which was significantly different from all the other methods of extraction. LgmPA was also significantly different from the other displaying a moderate underestimation.

Table 9 presents the PCA method of estimation for twenty variables and a sample size of one hundred. Again, the one-way analysis of variance procedure indicates that the null hypothesis was rejected. Therefore, we can conclude that there is at least one inequality among the methods of extraction. Tukey's multiple comparison procedure shows that the MAP00, MAP76, PA, and K1 are not significantly different. However, it is interesting to note that with the increase in sample size the MAP76 and K1 preformed perfectly under this scenario. Though Tukey's multiple comparison procedure showed no significant difference between MAP00 and PA from MAP76 and K1, there is something that should be noted. The MAP00 slightly overestimated the number of factors and the PA slightly underestimated the number of factors. This type of slight underestimation and

Table 8 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 20 N = 50)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	72.2716667	14.4543333	134.63	<.0001
Error	1794	192.6033333	0.1073597		
Corrected Total	1799	264.8750000			

R-Square	Coeff Var	Root MSE	NUM Mean
0.272852	10.95235	0.327658	2.991667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	72.27166667	14.45433333	134.63	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	TEST
	A	3.27333	300	SCREE
	B	3.05667	300	MAP00
	B			
	B	3.05000	300	MAP76
	B			
C	B	3.00000	300	K1
C				
C		2.97000	300	PA
	D	2.60000	300	LgmPA

overestimation should be taken into account by a researcher that might utilize these types of extraction. It is also interesting to note that Tukey's did not find a significant difference between PA and LgmPA. However, one should be aware of the fact that LgmPA tended to underestimate the factors more than PA. The SCREE and BART were significantly different from one another and all the other methods. While both methods overestimated the number of factors, the SCREE overestimated the number of factors to the point in which it was significantly different from BART's overestimation. It would seem that when the sample size is five times the number of variables such as it is in this scenario, the MAP76 and K1 tend to retain the true number of factors.

Table 9 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 20 N = 100)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	7.89142857	1.31523810	32.82	<.0001
Error	2093	83.87000000	0.04007167		
Corrected Total	2099	91.76142857			

R-Square	Coeff Var	Root MSE	NUM Mean
0.085999	6.619053	0.200179	3.024286

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	7.89142857	1.31523810	32.82	<.0001

Tukey's Studentized Range (HSD) Test for NUM

Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	TEST
	A	3.15000	300	SCREE
	B	3.07333	300	BART
	C	3.00333	300	MAP00
	C			
	C	3.00000	300	K1
	C			
	C	3.00000	300	MAP76
	C			
D	C	2.99333	300	PA
D				
D		2.95000	300	LgmPA

The PFA method of estimation for twenty variables with a sample size of one hundred is presented in Table 10. Since the null hypothesis was rejected in the one-way analysis of variance procedure, we can conclude that there is at least one inequality among the methods of extraction. Tukey's multiple comparison procedure shows that the MAP00, MAP76, PA, and K1 are not significantly different and performed perfectly in determining the number of factors to retain. This is a marked improvement for the extraction methods when compared to their results in Table 9 under the PCA estimation method. The SCREE was significantly different from all the other methods of extraction showing a tendency to overestimate. Also, the LgmPA was significantly different from all the other methods with a tendency to underestimate the true number of factors. Therefore, it can be concluded that the PFA estimation method has a better performance than PCA in methods of extraction for an added two more methods, MAP00 and PA.

Table 10 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 20 N = 100)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	22.7850000	4.5570000	89.75	<.0001
Error	1794	91.0900000	0.0507748		
Corrected Total	1799	113.8750000			

R-Square	Coeff Var	Root MSE	NUM Mean
0.200088	7.574207	0.225333	2.975000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	22.78500000	4.55700000	89.75	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.11000	300	SCREE
B	3.00000	300	K1
B			
B	3.00000	300	MAP00
B			
B	3.00000	300	MAP76
B			
B	3.00000	300	PA
C	2.74000	300	LgmPA

Tables 11 and 12 give the results from both estimation methods, PCA and PFA. The sample size was one hundred fifty for both estimation methods. It is interesting to note that the increase in sample size produced four methods of extraction that preformed perfectly in both estimation methods. Those methods of extraction are: K1, MAP00, MAP76, and PA. The SCREE was significantly different from all the other methods of extraction under both estimation methods. The SCREE overestimated in both scenarios. The LgmPA method of extraction only slightly under- estimated in the PCA method of estimation and was not found significantly different from the four that performed perfectly.

Table 11 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 20 N = 150)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	32.7790476	5.4631746	119.20	<.0001
Error	2093	95.9233333	0.0458305		
Corrected Total	2099	128.7023810			

R-Square	Coeff Var	Root MSE	NUM Mean
0.254689	7.008098	0.214081	3.054762

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	32.77904762	5.46317460	119.20	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	TEST
	A	3.35667	300	SCREE
	B	3.05000	300	BART
	B			
C	B	3.00000	300	K1
C	B			
C	B	3.00000	300	MAP00
C	B			
C	B	3.00000	300	MAP76
C	B			
C	B	3.00000	300	PA
C				
C		2.97667	300	LgmPA

BART's slightly overestimated, but it was not found significantly different from the four that preformed perfectly. However, BART and LgmPA were found to be significantly different from one another in the PCA method of extraction. Under the PFA method of extraction, the SCREE and LgmPA were found to be significantly different from one another and from the four that performed correctly in factor retention. The SCREE grossly overestimated while LgmPA underestimated. The substantive implication from this scenario is that the increase in sample size to a little over seven times the number of variables increased the precision of four methods of extraction for both methods of estimation.

Table 12 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 20 N = 150)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	83.7716667	16.7543333	216.89	<.0001
Error	1794	138.5833333	0.0772482		
Corrected Total	1799	222.3550000			

R-Square	Coeff Var	Root MSE	NUM Mean
0.376747	9.147636	0.277936	3.038333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	83.77166667	16.75433333	216.89	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.47667	300	SCREE
B	3.00000	300	K1
B			
B	3.00000	300	MAP00
B			
B	3.00000	300	MAP76
B			
B	3.00000	300	PA
C	2.75333	300	LgmPA

The PFA method of estimation for forty variables with a sample size of ninety is presented in Table 13. Since the null hypothesis was rejected in the one-way analysis of variance procedure, we can conclude that there is at least one inequality among the methods of extraction. Tukey's multiple comparison procedure shows that the MAP76, LgmPA, PA, and K1 are not significantly different. The K1, PA, LgmPA methods of extraction resulted in a perfect score in estimating the number of factors to retain. The MAP76 slightly overestimated the true number of factors. The SCREE, BART, MAP00 were significantly different from the other methods, but not from each other. All three methods of extraction tended to overestimate the true number of factors.

Table 13 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 40 N = 90)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	7.4295238	1.2382540	20.95	<.0001
Error	2093	123.7133333	0.0591081		
Corrected Total	2099	131.1428571			

R-Square	Coeff Var	Root MSE	NUM Mean
0.056652	7.952577	0.243122	3.057143

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	7.42952381	1.23825397	20.95	<.0001

Tukey's Studentized Range (HSD) Test for NUM

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.14000	300	SCREE
A			
A	3.12333	300	BART
A			
A	3.11000	300	MAP00
B	3.02667	300	MAP76
B			
B	3.00000	300	K1
B			
B	3.00000	300	PA
B			
B	3.00000	300	LgmPA

Table 14 is presented with the PFA estimation method under the same conditions as Table 13, according to the number of variables and sample size. The same three extractions methods performed perfectly in PFA as they did in the previous table with PCA. However, the Map00 was significantly different from all the other methods with a tendency to overestimate. The SCREE and MAP76 were not found to be significantly different. Both extraction methods tended to overestimate the true number of factors.

An interesting observation of these two tables compared to previous ones is the fact that LgmPA and PA performed much better. This observation tends to lend itself to the fact that the number of variables was doubled in these two tables. It is also interesting that both estimation methods performed the same for these two extraction methods. This convergence in methods of extraction was not seen in the case of twenty variables until the sample size was increased to one hundred variables.

Table 14 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 40 N = 90)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	3.10944444	0.62188889	19.29	<.0001
Error	1794	57.82333333	0.03223151		
Corrected Total	1799	60.93277778			

R-Square	Coeff Var	Root MSE	NUM Mean
0.051031	5.917533	0.179531	3.033889

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	3.10944444	0.62188889	19.29	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.11000	300	MAP00
B	3.06667	300	SCREE
C B	3.02667	300	MAP76
C	3.00000	300	K1
C	3.00000	300	PA
C	3.00000	300	LgmPA

Table 15 presents the estimation method PCA with forty variables with a sample size of two hundred. The BART method of extraction, which overestimated, was not found significantly different from the SCREE, which also overestimated. In contrast, Bart was significantly different from the others. The SCREE was not found significantly different from the others according to Tukey's multiple comparison procedure. However, the SCREE and MAP00 slightly overestimated while the rest (K1, MAP76, PA, and LgmPA) performed perfectly in determining the number of factors to retain. The same four methods

Table 15 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 40 N = 200)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.42285714	0.07047619	6.32	<.0001
Error	2093	23.34666667	0.01115464		
Corrected Total	2099	23.76952381			

R-Square	Coeff Var	Root MSE	NUM Mean
0.017790	3.508267	0.105616	3.010476

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	0.42285714	0.07047619	6.32	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	TEST
	A	3.040000	300	BART
	A			
B	A	3.020000	300	SCREE
B				
B		3.013333	300	MAP00
B				
B		3.000000	300	K1
B				
B		3.000000	300	MAP76
B				
B		3.000000	300	PA
B				
B		3.000000	300	LgmPA

of extraction performed perfectly in Table 16 with the PFA method of estimation. Table 16 also showed that the Map00 was not found significantly different from the perfect performers, but it did have the tendency to slightly overestimate. The SCREE, under PFA in Table 16, was significantly different from all the others with a tendency of overestimation. The two table's show that the increase in sample size to five times the number of variables help increase the accuracy of the methods of extraction. Similar results were seen in the case of twenty variables. Unlike the cases with twenty variables, we can now see that LgmPA and PA seem to respond better to a larger number of variables.

Table 16 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 40 N = 200)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	3.56277778	0.71255556	39.12	<.0001
Error	1794	32.67666667	0.01821442		
Corrected Total	1799	36.23944444			

R-Square	Coeff Var	Root MSE	NUM Mean
0.098312	4.468079	0.134961	3.020556

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	3.56277778	0.71255556	39.12	<.0001

Tukey's Studentized Range (HSD) Test for NUM
Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.12000	300	SCREE
B	3.00333	300	MAP00
B			
B	3.00000	300	K1
B			
B	3.00000	300	MAP76
B			
B	3.00000	300	PA
B			
B	3.00000	300	LgmPA

Table 17 presents the estimation method PCA with forty variables with a sample size of two hundred ninety. The only significant difference was the BART method of extraction, which overestimated. All the other methods of extraction were not found significantly different. Though they were not found significantly different, it must be noted that the SCREE, K1, and Map00 slightly overestimated. The MAP76, PA, and LgmPA all estimated the number of factors correctly. Again, it should be noted that LgmPA and PA perform better with the increase in variables and sample size.

Table 17 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PCA. (V = 40 N = 290)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.16476190	0.02746032	4.17	0.0004
Error	2093	13.76666667	0.00657748		
Corrected Total	2099	13.93142857			

R-Square	Coeff Var	Root MSE	NUM Mean
0.011827	2.698249	0.081102	3.005714

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	6	0.16476190	0.02746032	4.17	0.0004

Tukey's Studentized Range (HSD) Test for NUM

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.026667	300	BART
B	3.006667	300	SCREE
B			
B	3.003333	300	K1
B			
B	3.003333	300	MAP00
B			
B	3.000000	300	MAP76
B			
B	3.000000	300	PA
B			
B	3.000000	300	LgmPA

Table 18 presents the estimation method PFA with forty variables with a sample size of two hundred ninety. The results were very similar to the results from Table 17 using the PCA method of estimation. The only significant difference was the SCREE method of extraction, which overestimated. All the other methods of extraction were not found significantly different. Though they were not found significantly different, it must be noted that the K1 and Map00 slightly overestimated. The MAP76, PA, and LgmPA all estimated the number of factors correctly. Again, it should be noted that LgmPA and PA perform better with the increase in variables and sample size under both methods of estimation.

Table 18 – Analysis of Variance Procedure and Tukey's Multiple Comparison Procedure for PFA. (V = 40 N = 290)

The ANOVA Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	0.86444444	0.17288889	16.40	<.0001
Error	1794	18.91333333	0.01054255		
Corrected Total	1799	19.77777778			

R-Square	Coeff Var	Root MSE	NUM Mean
0.043708	3.409935	0.102677	3.011111

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEST	5	0.86444444	0.17288889	16.40	<.0001

Tukey's Studentized Range (HSD) Test for NUM

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	TEST
A	3.060000	300	SCREE
B	3.003333	300	K1
B			
B	3.003333	300	MAP00
B			
B	3.000000	300	MAP76
B			
B	3.000000	300	PA
B			
B	3.000000	300	LgmPA

It is evident from the Tables that the performance of the methods of extraction differed from each of the initial scenarios. Furthermore, it is interesting to note that as the sample size increased we see fewer significant differences among the methods of extraction. The increase in variables also created fewer significant differences among the methods of extraction. In the case with twenty variables, the MAP76, MAP00, PA, and K1 performed the best as the sample size increased. In the case of forty variables, the LgmPA, PA, and Map76 performed the best in determining the correct number of factors to retain. Both methods of estimation seem to parallel one another as the sample size was increased. This was evident in both the cases that involved twenty variables and forty variables.

The Mahalanobis distance was recorded for each of the six factor extraction methods when a change occurred in the number of factors retained that differed from the original population correlation matrix before the outlier was introduced. It should be noted that both estimation methods, PCA and PFA, were used in this process. The actual Mahalanobis distance that was recorded was the squared distance. After recording the squared distance for all six methods after a change in factor retention was detected, the data was used to construct ninety-five percent confidence intervals for each extraction method on each scenario. Tables 19-24 examine the performance of the six factor extraction methods with both estimation methods when an outlier is present in the data.

Table 19 examines the scenario of twenty variables with a sample size of fifty. All the methods of extraction over-estimated except for LgmPA, which underestimated 79% of the time. The mean squared Mahalanobis distance was similar for each method of estimation. The largest squared Mahalanobis distance that was recorded before the outlier was introduced, out of the original fifty population correlation matrices, was around thirty.

The methods of extraction did not start to deviate with the outlier until the squared Mahalanobis distance reached forty or more.

Table 19 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance. (V = 20 N = 50)

PCA				PFA			
TEST=K1							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
42.1721125	1.9881865	41.1126827	43.2315423	42.4151647	1.8819963	41.4475315	43.3827979
TEST=LgmPA							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
42.9713381	4.2413632	41.0406927	44.9019835	43.4730412	3.4799551	41.6838134	45.2622690
TEST=MAP00							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
41.1369727	3.3100707	39.9632727	42.3106727	41.1066697	3.2698615	39.9472272	42.2661122
TEST=MAP76							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
40.7104417	4.1269209	39.3140925	42.1067908	40.6826639	4.0945426	39.2972700	42.0680578
TEST=PA							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
43.2964682	3.0281063	41.9538811	44.6390553	43.5725188	2.3200743	42.3362384	44.8087991
TEST=SCREE							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
40.6780541	3.3999503	39.5444542	41.8116539	40.5523500	3.6606786	39.3137546	41.7909454

Table 20 has the same form as Table 19. The sample size was increased to one hundred in this scenario. It is interesting to note that PA was very resistant to outliers in this sample. However, all the other extraction methods performed similar to the previous Table. K1, MAP00, MAP76, and SCREE all overestimated. LgmPA was the only exception, which underestimated. The method of estimation, PCA and PFA, had similar results for all methods of extraction. The largest squared Mahalanobis distance that was recorded from the original fifty population correlation matrices was around thirty-one. The methods of extraction did not start to deviate with the presence of the outlier until the squared Mahalanobis distance reached seventy-two or more.

Table 20 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance. (V = 20 N = 100)

TEST	PCA Mean	PCA Std Dev	PCA Lower 95% CL for Mean	PCA Upper 95% CL for Mean	PFA Mean	PFA Std Dev	PFA Lower 95% CL for Mean	PFA Upper 95% CL for Mean
TEST=K1	81.7924615	3.1182689	79.9081105	83.6768125	81.0448700	3.1058501	78.8230786	83.2666614
TEST=LgmPA	84.1256800	9.0569006	72.8800526	95.3713074	72.7998429	19.2872707	54.9620926	90.6375931
TEST=MAP00	76.3525353	5.8060219	74.3267185	78.3783521	76.2568600	6.4165895	74.0526844	78.4610356
TEST=MAP76	74.9460257	6.1745541	72.8249922	77.0670592	74.8920778	6.7113355	72.6212885	77.1628670
TEST=PA	.	.	.	.	.	.	.	.
TEST=SCREE	74.0625762	12.3351868	70.2186614	77.9064910	74.0525872	11.1009518	70.4540749	77.6510995

Table 21 is presented in the same manner as Tables 19 and 20. In this case the sample size was increased to one hundred fifty. The largest squared Mahalanobis distance that was seen in the original data was thirty-five. The methods of estimation performed similar on all methods of extraction except for LgmPA. The LgmPA was more resistant to outliers under the PCA method of estimation. The change in factor retention, with the presence of the outlier, occurred when the squared Mahalanobis distance reached one hundred one or more.

Table 21 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance.
(V = 20 N = 150)

Test	PCA Mean	PCA Std Dev	PCA Lower 95% CL for Mean	PCA Upper 95% CL for Mean	PFA Mean	PFA Std Dev	PFA Lower 95% CL for Mean	PFA Upper 95% CL for Mean
TEST=K1	122.5058182	5.1305054	119.0590977	125.9525386	123.0020000	5.4165168	118.8384966	127.1655034
TEST=LgmPA	136.2480000	.	.	.	101.5653250	40.1602023	37.6614813	165.4691687
TEST=MAP00	111.4690419	11.7170924	107.1711777	115.7669061	111.1616226	11.4310551	106.9686777	115.3545674
TEST=MAP76	112.0317703	10.1304215	108.6541196	115.4094210	111.7742027	9.8701001	108.4833475	115.0650579
TEST=PA	132.3856000	6.7375166	124.0198683	140.7513317	126.9793333	4.7697642	123.3129680	130.6456987
TEST=SCREE	109.5997690	16.2981652	104.5209032	114.6786349	108.9671897	14.9848847	104.1096521	113.8247274

Tables 22–24 were all presented in the same manner as Tables 19-21. However, these tables had forty variables with different sample sizes. The largest squared Mahalanobis distance seen in the original data for the following tables was fifty-five, fifty-six, and sixty-eight. The change in factor retention, with an outlier present occurred when the squared Mahalanobis distance for the following tables reached around seventy-two, one hundred forty-five, and two hundred. The mean distance for all methods of extraction was very similar for each estimation method. All methods of extraction overestimated with the presence of the outlier except for LgmPA, which underestimated. However, it should be noted that when the sample size was two hundred, LgmPA was very resistant to the outlier in both PCA and PFA. Also, in the PCA estimation method, the PA method of extraction was resistant to outliers with a sample size of two hundred.

To further summarize the data with an outlier, box plots were created. Each box plot shows the squared Mahalanobis distance for each scenario used in the study. Figures 1-12 in the appendix presents these box plots.

After studying the squared Mahalanobis distance of data containing outliers, a very similar pattern emerged from the data. That pattern was discovered in utilizing the lower-bound of the confidence interval with the smallest squared Mahalanobis distance for each Table excluding Longman's Parallel Analysis. This lower bound was then compared to the maximum squared Mahalanobis distance from the original population correlation matrix. The patterns that emerged are as follows for the various sample sizes. It should also be noted that the patterns held for each scenario with a different number of variables. When the sample size is 2V+10, the methods of extraction were not affected until the outlier was at least 1.27 times greater than the largest squared Mahalanobis distance from the original data sets. With a sample size of 5V, the methods of extraction were not affected until the outlier was at least 2.25 times greater than the largest squared Mahalanobis distance from the original data sets. A sample size of 7V+10 did not show an effect until the outlier was at least 2.85 times greater than the largest observed squared Mahalanobis distance from the original data sets. Longman's Parallel Analysis was excluded due to erratic behavior that can be seen in the Tables.

Table 22 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance. (V = 40 N = 90)

PCA				PFA			
---------- TEST=K1 ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
78.8192395	3.8807521	77.6249203	80.0135588	79.4171651	3.4988587	78.3403753	80.4939549
---------- TEST=LgmPA ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
85.0901500	0.7455027	78.3920742	91.7882258	84.1898333	1.6460855	80.1007303	88.2789364
---------- TEST=MAP00 ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
73.4891184	6.2405507	71.6966229	75.2816138	73.5303880	6.1834333	71.7730757	75.2877003
---------- TEST=MAP76 ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
74.3756580	5.5305291	72.8038990	75.9474170	74.3756580	5.5305291	72.8038990	75.9474170
---------- TEST=PA ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
84.0789333	1.4591909	82.5476064	85.6102602	82.9714938	2.5794083	81.5970241	84.3459634
---------- TEST=SCREE ----------							
Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
72.7681409	7.1475623	70.5950834	74.9411985	73.3853667	7.2658615	71.2024602	75.5682731

Table 23 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance.
(V = 40 N = 200)

Test	PCA Mean	PCA Std Dev	PCA Lower 95% CL for Mean	PCA Upper 95% CL for Mean	PFA Mean	PFA Std Dev	PFA Lower 95% CL for Mean	PFA Upper 95% CL for Mean
TEST=K1	166.2705938	8.3587723	163.2569350	169.2842525	165.2121429	7.9366971	162.1346130	168.2896727
TEST=LgmPA	.	.	.	.	.	.	.	.
TEST=MAP00	145.8887083	12.0840837	142.3798562	149.3975605	145.8887083	12.0840837	142.3798562	149.3975605
TEST=MAP76	146.7813673	12.8826439	143.0810400	150.4816947	146.7813673	12.8826439	143.0810400	150.4816947
TEST=PA	.	.	.	.	165.4815000	7.8845566	162.2968587	168.6661413
TEST=SCREE	150.2876667	11.0590571	146.8414219	153.7339115	150.2195500	13.7986501	145.8065276	154.6325724

Table 24 – 95% Confidence Intervals for the Mean Squared Mahalanobis Distance.
(V = 40 N = 290)

TEST	PCA Mean	PCA Std Dev	PCA Lower 95% CL for Mean	PCA Upper 95% CL for Mean	PFA Mean	PFA Std Dev	PFA Lower 95% CL for Mean	PFA Upper 95% CL for Mean
TEST=K1	249.8433667	15.2821361	244.1369233	255.5498101	252.2929231	14.4938876	246.4387152	258.1471310
TEST=LgmPA	256.8230000	7.2572578	238.7949722	274.8510278	257.9710000	11.7195878	152.6746814	363.2673186
TEST=MAP00	201.1798298	18.6658187	195.6993401	206.6603195	200.3074783	19.4235162	194.5394060	206.0755506
TEST=MAP76	200.9311042	16.1468379	196.2425512	205.6196572	200.0720213	16.9298094	195.1012430	205.0427995
TEST=PA	259.2714286	6.3705447	253.3796574	265.1631998	252.6303600	14.6881586	246.5673861	258.6933339
TEST=SCREE	215.2868913	22.8873519	208.4901868	222.0835959	210.4950682	29.6769668	201.4724456	219.5176908

It should be noted that the squared Mahalanobis distance was computed for each of the fifty original population correlation matrices that were randomly selected from each scenario containing three hundred population correlation matrices. It is of interest that the maximum squared Mahalanobis distance for each population correlation matrix had at least two or three squared Mahalanobis distances very close to the maximum squared

Mahalanobis distance. This was true for each scenario, regardless of the sample size or the number of variables.

CHAPTER 4: SUMMARY

Summary of Phase One

The first phase of this empirical study, involving the methods of factor extraction with multivariate normal data, clearly suggests that the choice of method and the design of the factor analytic study, according to the number of variables and sample size, play a crucial role in determining the correct number of factors to retain. From the data, it was evident that larger sample sizes affected the methods of extraction in retaining the correct number of factors. It should also be noted that even with a smaller number of variables it was still evident that the larger sample sizes played a crucial role for certain methods of extraction.

In terms of overall accuracy, the K1 and PA methods of extraction provided the largest portion of correct decisions in retaining factors. However, it should be noted that data sets consisting of twenty variables showed better performances in determining the correct number of factors to retain when the sample size was increased. Furthermore, in the cases involving twenty variables, a sample size of five times the number of variables yielded the perfect score in retaining the correct number of factors for K1 and MAP76, in the Principal Component Analysis (PCA) estimation method. The same number of variables and sample size in the Principal Factor Analysis (PFA) estimation method had K1, MAP00, MAP76, and PA with a perfect score in retaining the correct number of factors. When the sample size was increased to more than seven times the number of variables, both, Principal Component Analysis (PCA) and Principal Factor Analysis (PFA), yielded the same four extraction methods with perfect retention of the true factors. The four methods of extraction are as follows: K1, MAP00, MAP76, and PA.

The increase in variables in the study brought different results for which method would retain the correct number of factors. LgmPA and PA had the best results for both estimation methods. In fact, both methods of extraction correctly identified the true number of factor to retain under each scenario involving forty variables. The Map76 correctly retained the number of factors when the sample size was greater than or equal to five times the number of variables for both estimation methods. The K1 method began to overestimate as the sample size was increased to more than seven times the number of variables.

In concluding the first phase of this empirical study, it is recommended that researchers should try to conduct their research with a sample size of a little more than seven times the number of variables when possible. Parallel Analysis would be the best overall recommended method of extraction to be utilized in both Principal Component Analysis and Principal Factor Analysis with sample sizes of this nature.

Summary of Phase Two

The second phase of this empirical study investigated which method of extraction in factor analysis is least resistant to outliers, when they are present in the data. The research indicates that Parallel Analysis and Longman's Parallel Analysis was very resistant to outliers in some specific cases. However, it was evident from the data that each method tended to make the incorrect decision on retaining the correct number of factors when the squared Mahalanobis distance reached a certain amount. Therefore, as a rule of thumb, a researcher might want to calculate the squared Mahalanobis distance for all points to the mean to find the actual distance that the points lie from the mean. If the squared Mahalanobis distance of the furthest point is beyond a certain amount from the next highest squared Mahalanobis distance, then the researcher needs to choose an appropriate method of extraction for determining the number of factors. The theory behind this statement comes from an examination of the original population correlation matrices in which several squared Mahalanobis distances were noted as being very close to the maximum squared Mahalanobis distance. The amount of difference is subject to the sample size of the study at hand. For example, using the pattern that emerged from this empirical study, a researcher would not have reason for alarm until the following conditions were experienced. When the sample size is 2V+10, the methods of extraction will not be affected until the outlier is at least 1.27 times greater than the next largest squared Mahalanobis distance from the data set. With a sample size of 5V, the methods of extraction will not be affected until the outlier is at least 2.25 times greater than the next largest squared Mahalanobis distance from the data set. A sample size of 7V+10 will not show an effect until the outlier is at least 2.85 times greater than the next largest observed squared Mahalanobis distance from the data set.

Additional work will be necessary to fully explore the nature of the outlier in exploratory factor analysis. The importance of this additional work is evident when one just realizes the extent in which this multivariate technique is used across all disciplines. Needless to say, that this multivariate technique is frequently used in educational research. However, it was the intention of this researcher to explore this matter due to the fact that in the real world of data collection and analysis no data set is perfectly normal or multivariate normal. It stands to reason that researchers that do encounter a potential outlier or actual outlier need a way of assessing and addressing them.

Hopefully, this study will be informative to researchers who are conducting studies that deal with data reduction or detecting data structure. One of the main reasons that this study will be informative to researchers is due to the fact that real data are never exactly multivariate normal (Johnson & Wichern, 1988). The question that faces many researchers is how many factors to retain in factor analysis. The interpretations of the data rely in large part upon the extraction of the correct number of factors. Therefore, the methods of factor extraction play a crucial role and researchers must be aware of the limitations of certain methods and procedures. The results of this empirical study should highlight the need for researchers to exercise caution in the methods of factor extraction. Also, the planning stages of their research should be carefully considered.

Implications for Practice

The second phase of the empirical study dealt with the affect of an outlier on the methods of extraction. The importance of this research is to allow the researcher to identify the squared Mahalanobis distance of a potential outlier in a factor analytical study that will affect the methods of extraction. Outliers can have a profound effect on the analysis of a study. The outlier has the potential to distort the variance and covariance of the data (Moser, 2004). If the researcher is aware of the presence of a potential outlier, knowing the squared Mahalanobis distance in which the methods of extraction start to fail in identifying the true number of factors could prove invaluable. Even though the covariance and variance structure could be affected with the presence of the outlier, the researcher will be able to actually know the true number of factors if the squared Mahalanobis distance is not beyond a certain amount. This will aid the researcher in interpretation of the underlying factors by not having to deal with issues of overestimation or underestimation.

In an Exploratory Factor Analysis, the decision to retain the correct number of factors is crucial because it is made prior to factor rotation. As mentioned earlier in the paper, Barry Moser (2004) mentions that factor rotation methods are utilized to find equivalent solutions that are easier to interpret. Therefore, knowledge of the true number of factors, will aid the researcher due to the fact that a true Exploratory Factor Analysis is one which is conducted in which the researcher has no true idea of the number of factors that underlie the data.

To deal with the issue of an outlier once it is detected in the analysis, the researcher has to decide what options are available on how to deal with the issue. One such option that the researcher has is a Sensitivity Analysis. An analysis of this source allows the researcher to assess the relative importance of model input factors (Saltelli, Tarantola, Campolongo, & Ratto, 2004). A Sensitivity Analysis, in practice, would have the researcher remove the outlier from the data and repeat the factor analysis without the outlier and compare the results for major changes (Moser, 2004). Another option, the researcher might want to consider, is robust methods for estimating the covariance matrix. These robust methods of estimating the covariance matrix works on transforming the data prior to entering it into the factor analysis. In current research, some suggest utilizing an isometric log-ratio transformation (Egozcue, Pawlowsky-Glahn, Mateu-Figueraz, & Barcel´o-Vidal, 2003). However, other researchers suggest that using the isometric log-ratio transformation to obtain a robust estimation of the covariance matrix can lead to some uninterruptable results (Filzmoser, Hron, Reimann, & Garrett, 2009). Therefore, Filzmoser et al. (2009) feel a back transformation of the isometric log-ratio results to a centred log-ratio transformation will allow better interpretations for the researcher.

Limitations

The results of this research must be taken in the light of the limitations that this empirical study exhibits. Although the simulation approach followed has examined a range of values for the number of variables and sample size, further research is recommended to extend the findings in this empirical study. Perhaps, additional work is needed with population correlation matrices that have more underlying factors. Outliers could be incorporated in the same manner as in this empirical study to see if the rule of thumb actually holds true under a different number of factors.

The level of factor saturation in this study was set to values ranging from 0.2 to 0.8. It has been shown that ranges from 0.4 to 0.8 are wide enough to show a difference in decision rules for methods of extraction (Linn, 1968). Zwick and Velicer (1982) felt the need to set the lower saturation level at 0.5, in order to avoid trivial loadings, and the upper level at 0.8. It has been noted in research studies that loadings of 0.85 and above are rarely found in practice (Zwick & Velicer, 1982). However, when dealing with real data, the researcher never knows what he/she will actually obtain. Therefore, the population correlation matrices could have been simulated to exhibit different levels of saturation. For example, the level of factor saturation could have been set to values ranging from 0.6 to 0.8 to represent a high level of factor saturation. Furthermore, a low level of factor saturation with values ranging from 0.2 to 0.4 could have been employed into the study.

REFERENCES

Archambault, I., Janosz, M., Fallu, J.-S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence , 32* (3), 651 - 670.

Bartlett, M. S. (1950). Test of significance in factor analysis. *British Journal of Psychology , 3*, 77-85.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology , 4*, 1-2.

Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (fifth edition ed.). White Plains, NY: Longman.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* New York: Plenum Press.

Child, D. (1978). *The essentials of factor analysis.* London: Holt, Rinehart and Winston, Inc.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.

Egozcue, J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology , 35* (3), 279-300.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods , 4* (3), 272-299.

Fan, X., Felsovalyi, A., Sivo, S., & Keenan, S. (2001). *SAS for monte carlo studies: A guide for quantitative researchers.* Cary, NC: SAS Institute Inc.

Fava, J., & Velicer, W. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research , 27* (3), 387-415.

Fava, J., & Velicer, W. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research , 27* (3), 387-415.

Filzmoser, P., Hron, K., Reimann, C., & Garrett, R. (2009). Robust factor analysis for compositional data. *Computers & Geosciences , 35* (9), 1854-1861.

Franklin, S., Gibson, D., Robertson, P., Pohlmann, J., & Fralish, J. (1995). Parallel analysis: A method for determining significant principal components. *Journal of Vegetation Science , 6* (1), 99-106.

Garson, G. D. (1998). "Factor Analysis", from *Statnotes: Topics in multivariate analysis.* Retrieved Jan. 31, 2007, from http://www2.chass.ncsu.edu/garson/pa765/factor.htm

Glorfeld, L. W. (1995). An improvement on Horn's Parallel Analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement , 55* (3), 377-393.

Gorsuch, R. L. (1974). *Factor analysis*. London: W. B. Saunders.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). New Jersey: Lawrence Erlbaum Associates, Inc.

Hatcher, L. (2007). *A step-by-step approach to using SAS® for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Inc.

Hatcher, L. & Stepanski, E. J. (1994). *A step-by-step approach to using the SAS system for univariate and multivariate statistics*. Cary, NC: SAS Institute Inc.

Harmen, Harry H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

Hetzel, R.D. (1995, April). *The use of factor analysis in counseling psychology research: Current practice is not necessarily best practice*. Paper presented at the annual meeting of the American Education Research Association, San Francisco, CA.

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

Johnson, R. A. & Wichern, D. W. (1988). *Applied multivariate statistical analysis* (2nd ed.). New Jersey: Prentice-Hall Inc.

Kerlinger, F. N. & Pedhazur, E. J. (1968). Educational attitudes and perceptions of desirable traits of teachers. *American Educational Research Journal,5* (4), 543-560.

Kim, J.-O., & Mueller, C.W. (1978). *Factor analysis: Statistical methods and practical issues.* London: Sage Publications, Inc.

Kim, J.-O., & Mueller, C.W. (1978). *Introduction to factor analysis: What it is and how to do it.* London: Sage Publications, Inc.

Kottkamp, R.B., Mulhern, J.A., & Hoy, W.K. (1987). Secondary school climate: A revision of the OCDQ. *Educational Administration Quarterly, 23* (3), 31-48.

Linn, R. L. (1968). A monte carlo approach to the number of factors problem. *Psychometrika, 33* (1), 37-71.

Longman, S., Cota, A., Holden, R., & Fekken, C. (1989). A regression equation for the parallel analysis criterion in principal components analysis: Mean and 95th percentile eigenvalues. *Multivariate Behavioral Research, 24* (1), 59-69.

Loonstra, B., Brouwers, A., & Tomic, W. (2009). Feelings of existential fulfilment and burnout among secondary school teachers. *Teaching and Teacher Education: An International Journal of Research and Studies, 25* (5), 752-757. (ERIC Document Reproduction Service No. EJ840181) Retrieved July 29, 2009, from ERIC database.

Messick, S. (1990). Meaning and values in test validation: The science and ethics of test assessment. *Educational Researcher*, *18* (2), 5-11.

Moser, E. B. (2004). *Multivariate statistical data analysis*. Unpublished manuscript, Louisiana State University at Baton Rouge.

Rencher, A. C. (1995). *Methods of multivariate analysis*. New York: John Wiley & Sons, Inc.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. West Sussex, England: John Wiley & Sons, Ltd.

SAS Institute Inc. (1985). *SAS/IML user's guide, version 5 edition*. Cary, NC: Author.

SAS Institute Inc. (2004). SAS 9.1.3 Help and Documentation: Your Complete Guide to Syntax, How To, Examples, Procedures, and Concepts, What's New, and Tutorials. Version 5 Edition. Cary, NC: Author, 2002-2004.

Skinner, H. A. (1979). Dimensions and clusters: A hybrid approach to classification. *Applied Psychological Measurement, 3* (3), 327-388.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A monte carlo study. *Psychological Bulletin*, *106* (1), 148-154.

StatSoft (2008). Electronic Statistics Textbook is a public service provided by StatSoft, Inc. Retrieved March 8, 2008, from http://www.statsoft.com/textbook/glosfra.html

Stevens, J. (1986). *Applied multivariate statistics for the social sciences*. New Jersey: Lawrence Erlbaum Associates Inc.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Tanguma, J. (1999). *Determining the number of factors to retain*. Paper presented at the annual meeting of the Education Research Exchange, Texas A & M University, College Station, TX.

Thorndike, R. M. (1997). *Measurement and evaluation in psychology and education* (6^{th} ed.). New Jersey: Prentice-Hall, Inc.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology, 34* (4), 414-424.

Tucker, L. B., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34* (4), 421-459.

Turner, N.E. (1998). The effect of common variance and structure patterns on random data eigenvalues: Implications for accuracy of parallel analysis. *Educational and Psychological Measurement, 58* (4), 541-568.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. pp. 41-71 in R. D. Goffin and E. Helmes, eds., Problems and solutions in human assessment Boston: Kluwer.

Wilkinson, L., G. Blank, and C. Gruber (1996). *Desktop data analysis with SYSTAT*. New Jersey: Prentice-Hall Inc.

Zwick, W. R. & Velicer, W. F, (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17* (2), 253-269.

Zwick, W. R. & Velicer, W. F, (1984, August). *A Comparison of Five Rules for Determining the Number of Components in Data Sets*. Paper presented at the Annual Meeting of American Psychological Association, Toronto, Ontario.

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

APPENDIX: BOXPLOTS

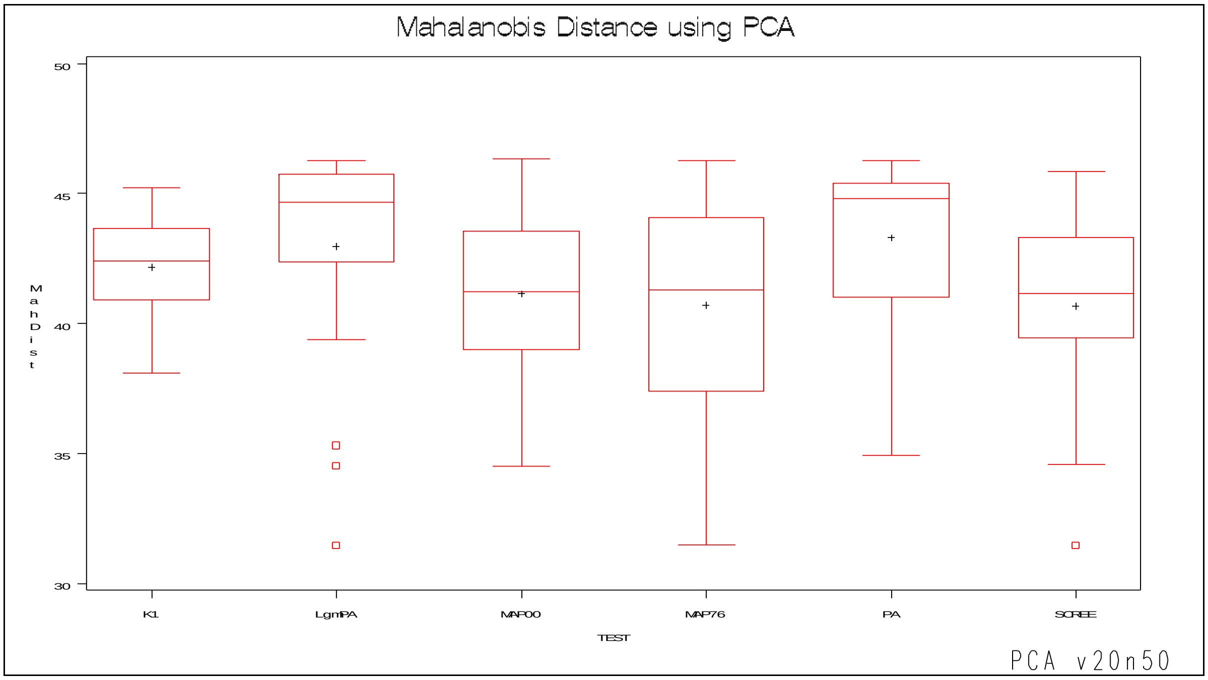

Figure 1 – PCA (V = 20 N = 50)

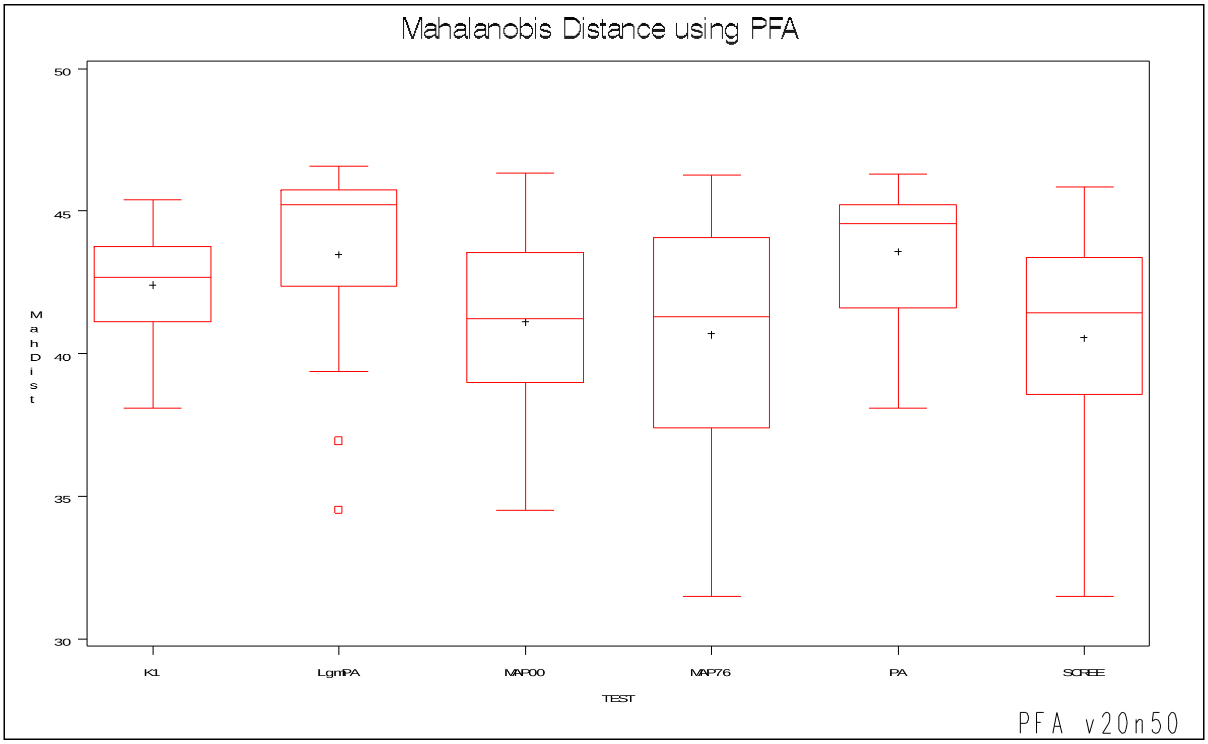

Figure 2 – PFA (V = 20 N = 50)

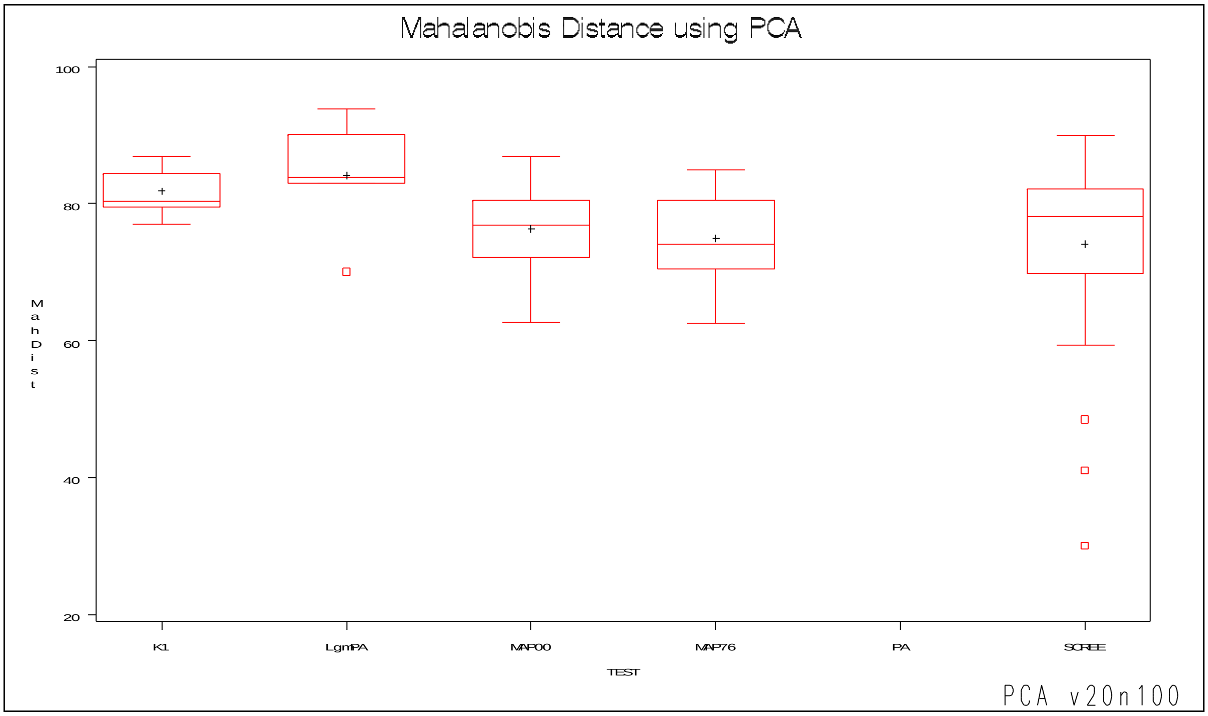

Figure 3 – PCA (V = 20 N = 100)

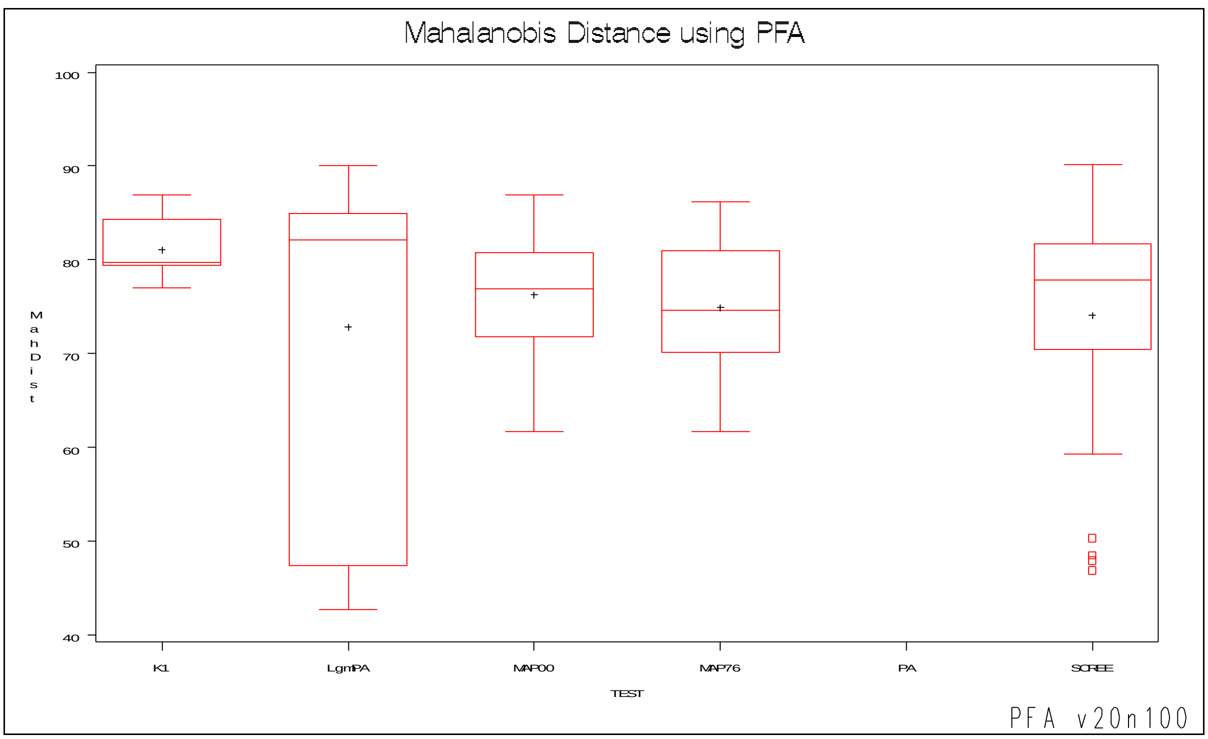

Figure 4 – PFA (V = 20 N = 100)

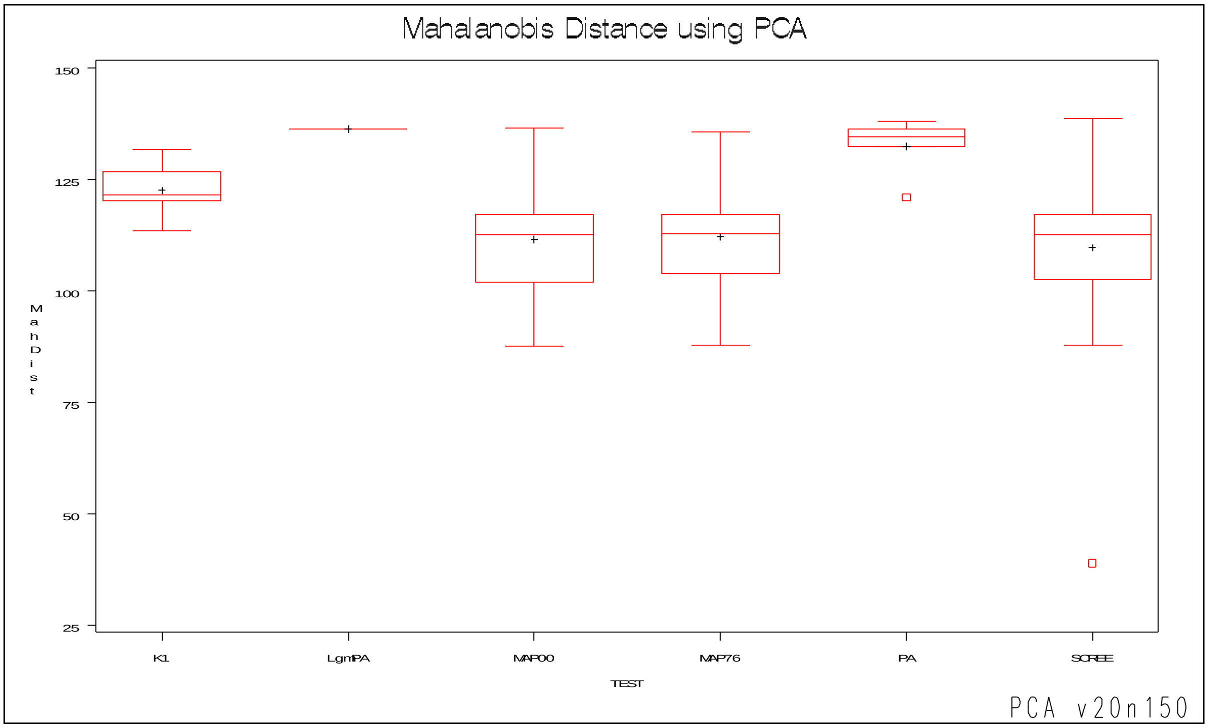

Figure 5 – PCA (V = 20 N = 150)

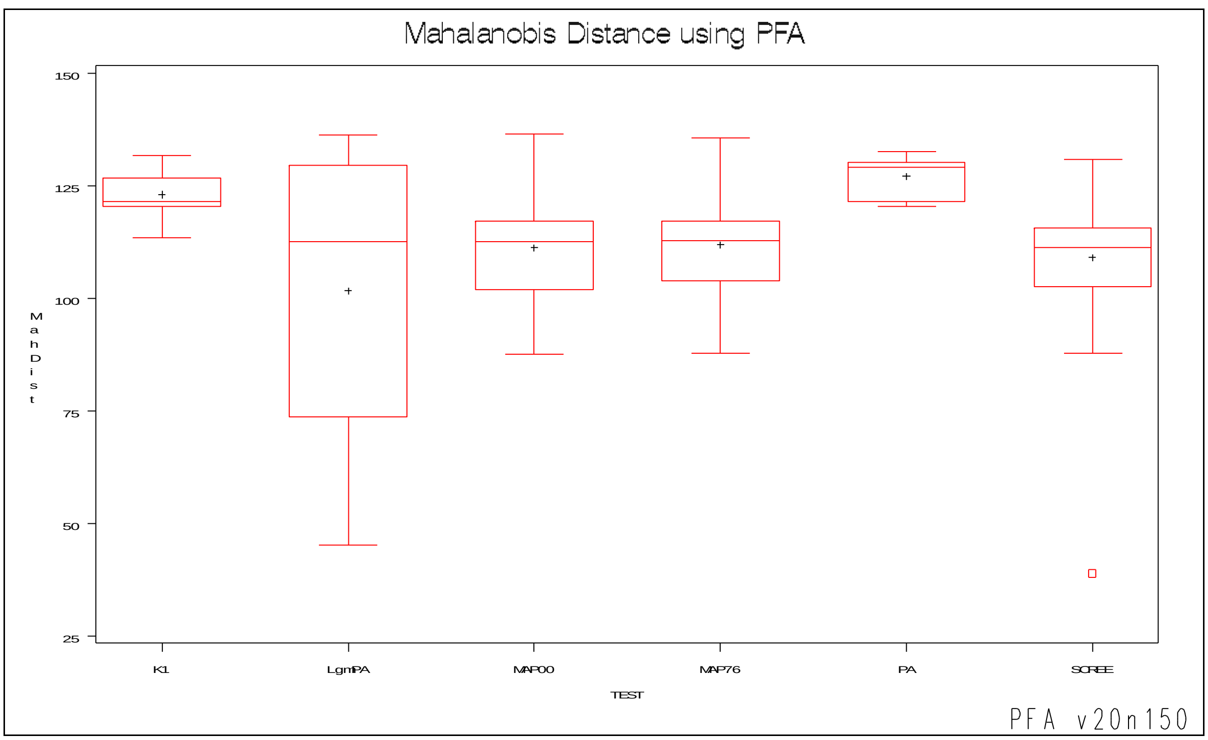

Figure 6 – PFA (V = 20 N = 150)

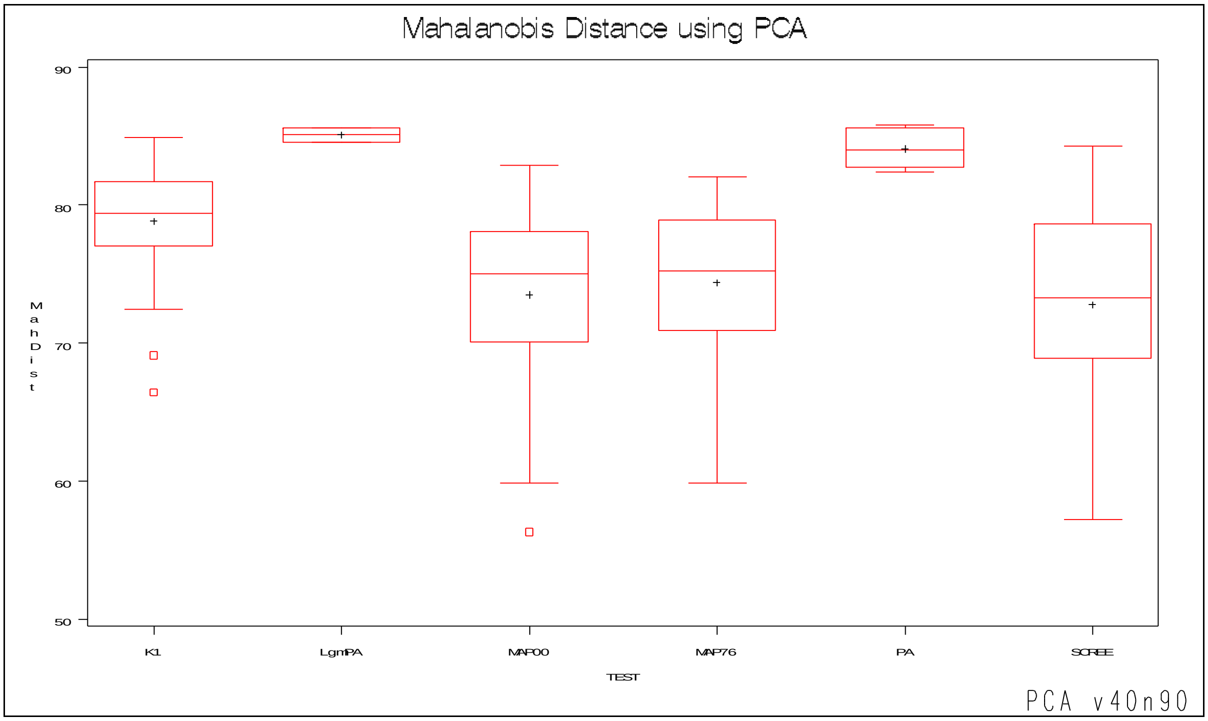

Figure 7 – PCA (V = 40 N = 90)

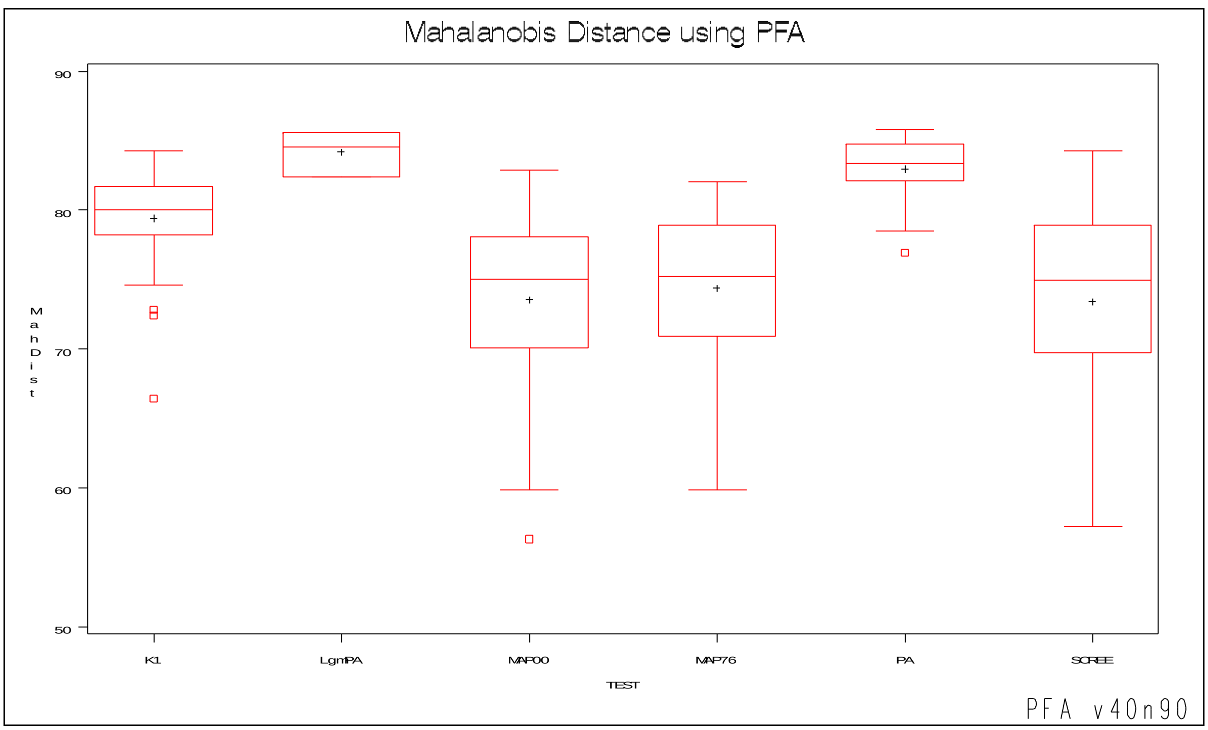

Figure 8 – PFA (V = 40 N = 90)

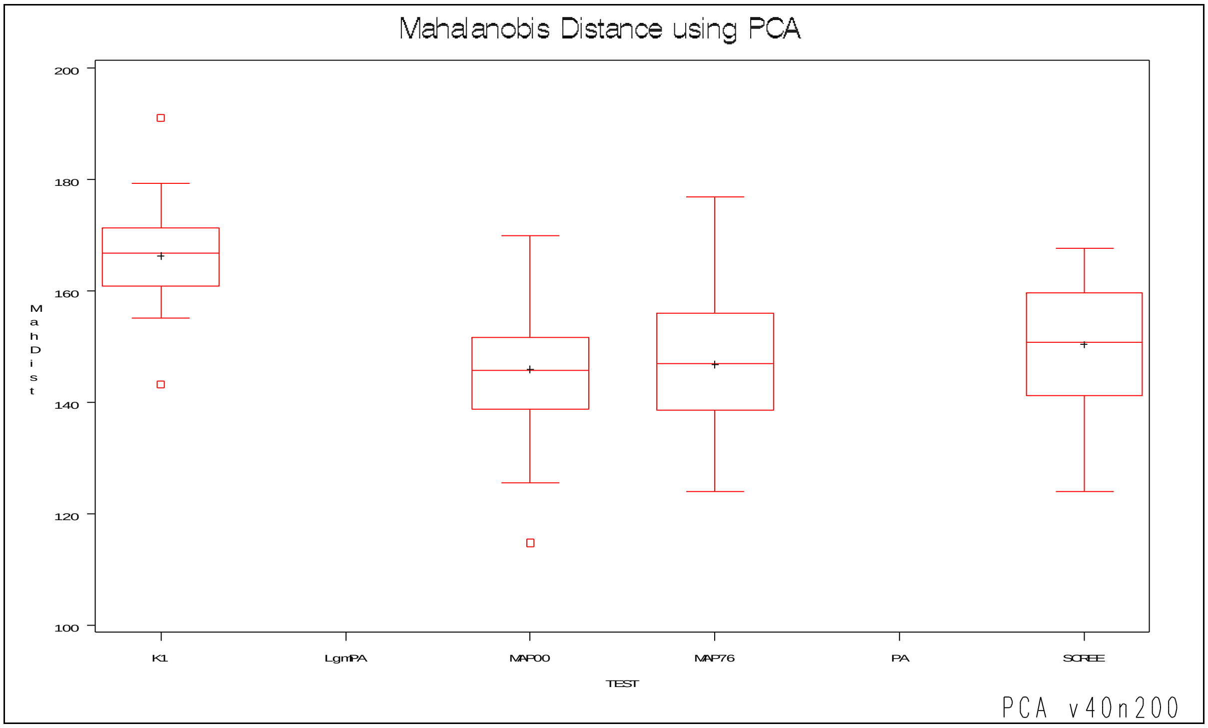

Figure 9 – PCA (V = 40 N = 200)

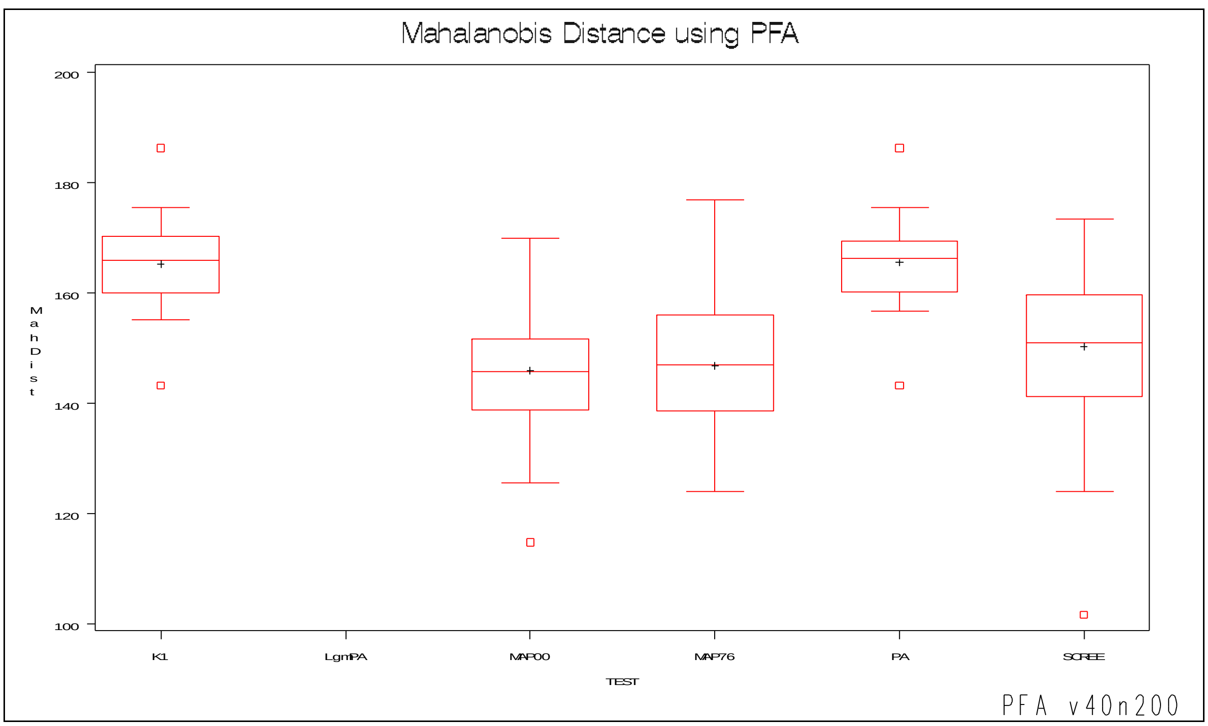

Figure 10 – PFA (V = 40 N = 200)

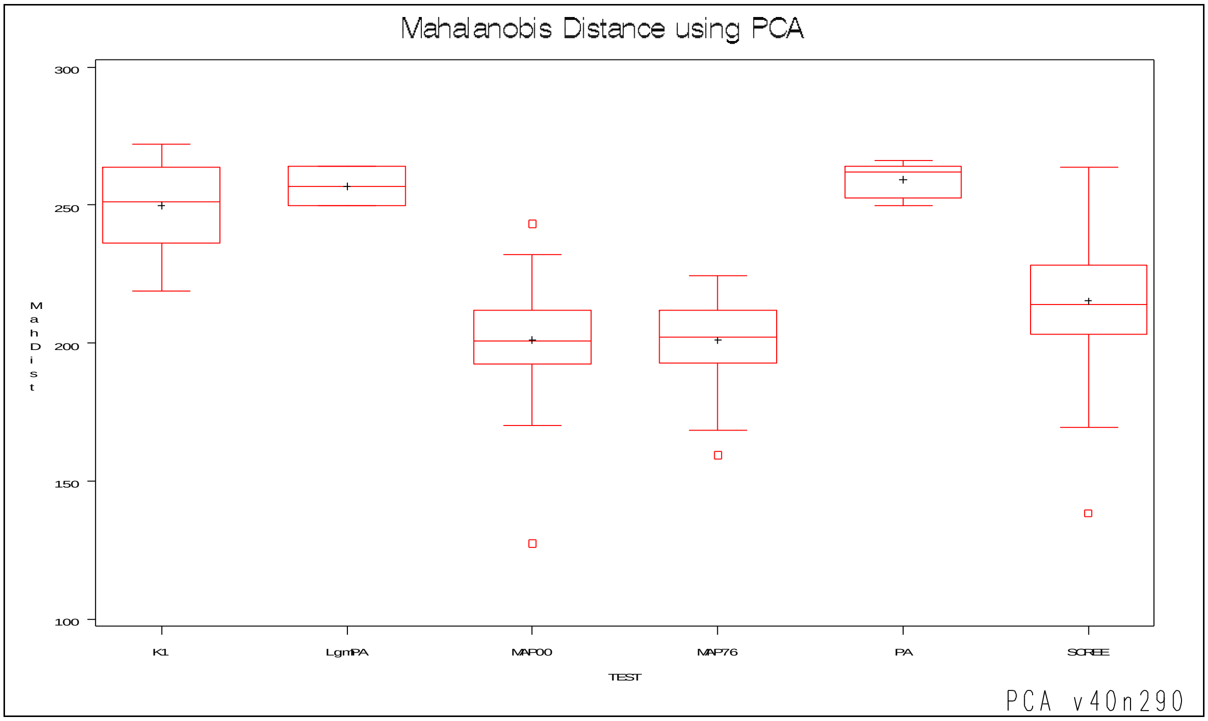

Figure 11 – PCA (V = 40 N = 290)

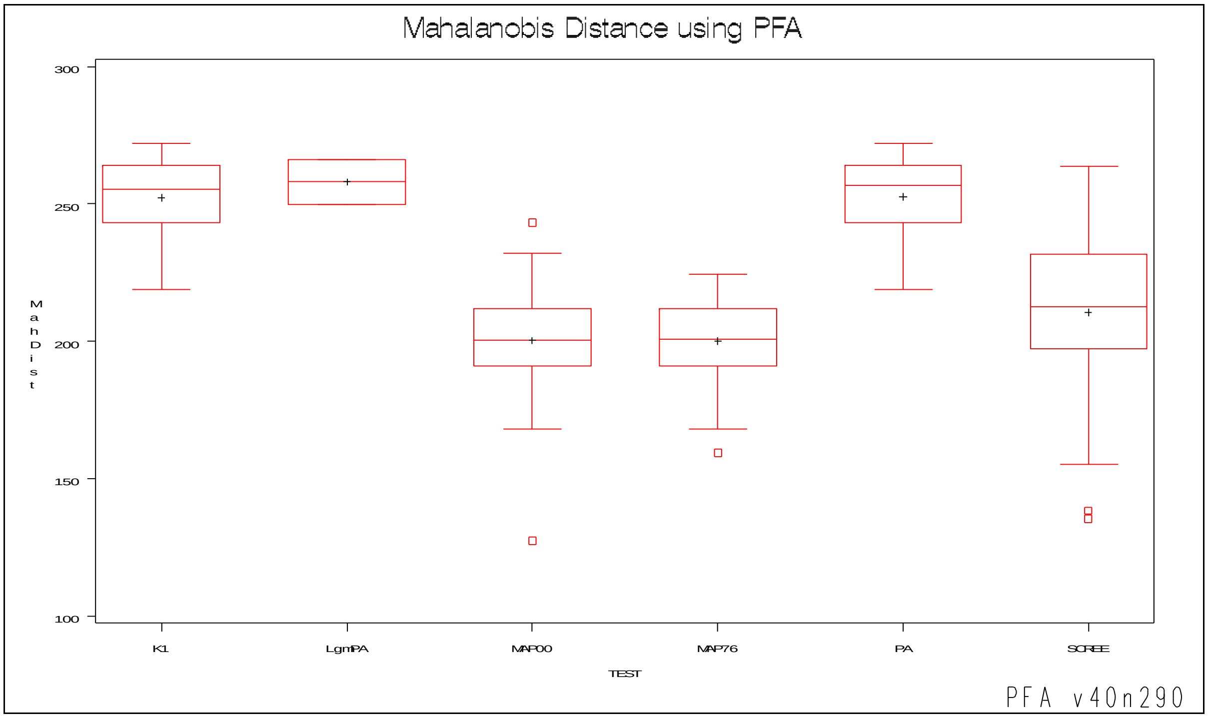

Figure 12 – PFA (V = 40 N = 290)

VITA

Victor Snipes Swaim was born November, 1968, in Memphis, Tennessee, to Olivia Jeanette Swaim and the late Harry Edward Swaim. He graduated from Hammond High School in 1986. He received a Bachelor of Science degree, majoring in mathematics, from Middle Tennessee State University. In 1992, Mr. Swaim returned to Louisiana to pursue a master's degree in applied statistics. During the time he was attending graduate school, he accepted a job teaching mathematics and coaching football at his high school alma mater, Hammond High School. The next year Mr. Swaim accepted a job teaching mathematics at Saint Thomas Aquinas High School. He was instrumental in helping start the high school football program under his old coach Pete Valenti. During this time, Mr. Swaim graduated from Louisiana State University with a master's in applied statistics in 1997.

From 1998-2004 Mr. Swaim taught at Albany High School where he was the head football coach until 2002. During the time he worked for the Livingston Parish School Board, he completed the course work at Louisiana State University in the Department of Educational Theory, Policy, and Practice to become certified as a Provisional Secondary School Principal (6-12), Provisional Principal (k-12), and a Supervisor of Student Teaching (1-12). Also, during this time he began to work as a part-time instructor at Southeastern Louisiana University in the Department of Mathematics. He taught business calculus, college algebra, and elementary statistics for the university. Working at Southeastern was a career defining experience that motivated Mr. Swaim to return to graduate school to obtain his doctoral degree in Educational Leadership and Research specializing in statistics.

In 2004, Mr. Swaim was hired as a full-time instructor in the Department of Mathematics at Southeastern Louisiana University. Presently, he is still employed at Southeastern Louisiana University as a permanent full-time instructor. He serves on several departmental committees and is active in campus activities.