(1)

A Monte Carlo Study Of The Raw And Normal Varimax Rotation Criterion In Factor Analysis

THESIS

William Michael Ibinson
Captain, USAF

AFIT/GOR/ENS/93M-27

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GOR/ENS/93M-27

A Monte Carlo Study Of The Raw And Normal
Varimax Rotation Criterion In
Factor Analysis

THESIS

William Michael Ibinson
Captain, USAF

AFIT/GOR/ENS/93M-27

AFIT/GOR/ENS/93M-27

A Monte Carlo Study Of The Raw And Normal Varimax Rotation Criterion In Factor Analysis

THESIS

Presented to the Faculty of the School of Operational Sciences
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Masters of Science in Operations Research

DTIC QUALITY INSPECTED 4

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

William Michael Ibinson, B.S.
Captain, USAF

23 March, 1993

Approved for public release; distribution unlimited

THESIS APPROVAL

STUDENT: Capt William M. Ibinson CLASS: GOR-93M

THESIS TITLE: A Monte Carlo Study of Raw and Normal Varimax Rotation Criterion in Factor Analysis

DEFENSE DATE: 5 March, 1993

COMMITTEE:	NAME/DEPARTMENT	SIGNATURE
Advisor	LtC Kenneth Bauer/ENS	Kenneth W Bauer Jr.
Reader	Dr Joseph P. Cain/ENS	Joseph P Cain

Preface

This study uses a Monte Carlo approach to address rotation techniques in a multivariate statistical technique known as factor analysis. This effort was sponsored by the Department of Operational Sciences, School of Engineering, Air Force Institute of Technology (AU) at Wright-Patterson AFB, OH.

I would like to thank the faculty and staff of the Air Force Institute of Technology, particularly my reader, Dr Joseph P. Cain. I wish to express my sincere appreciation to my advisor, Lt. Col. Keneth Bauer, Whose insight, patience, and damage control kept this effort on track.

Also thanks to capt s Dave Legge, Cindy Brown, Jeff lanning and Lt Mike Silver for their support and friendship.

Finally, I wish to thank my wife and daughter for their love and understanding during this program. It was truly a team effort.

Remember, God created time and no man can prolong his allotted hours. He can only live them to the fullest.

William Michael Ibinson

Table of Contents

Page

List of Figures

List of Tables

Table | Page

Abstract

Factor analysis is a multivariate statistical procedure for analyzing and reducing large data sets. Many factor analysis schemes and techniques are available that lead to strikingly different results from the same data. This research effort used a Monte Carlo approach to investigate the properties of two rotation methods for simple structure, Kaiser's raw and normal varimax criterion. Data sets were developed from a set of contrived experimental factor patterns by multiplying each factor pattern by its transpose to create a covariance matrix. Data sets of multivariate normal deviates were in turn generated from each covariance matrix via the Choleski algorithm. Rotated factor pattern matrices from each data set were compared to their respective experimental factor pattern on the basis of structure, loadings and eigenvalues. These performance issues are addressed through regression analysis and separate factor analysis in which the grand mean of proposed measures of effectiveness are predicted. These measures of effectiveness include structure matching and root mean square error between the experimental and observed factor patterns. Several methods of characterizing factor pattern complexity and predicting rotation criterion performance are explored.

A Monte Carlo Study Of The Raw And Normal Varimax Rotation Criterion In Factor Analysis

I. Introduction

The concept that nations must be prepared to defend themselves at all times from every type of outside threat is a relatively modern one. Prior to World War II, countries relied on troop mobilization and conscription rather than a large standing army. Fortunate for the ill-equipped and poorly trained troops entering WW II, our scientific and industrial communities were able to put new technology on the battle field in numbers large enough to bring about a favorable end. Also fortunate was the successful application of quantitative economic analysis and other scientific methods to military operations. Since that time, we have seen exponential increases not only in weapon lethality but also in the speed and range of weapon deployment. The time to develop and transition new technology to the battlefield is a luxury we no longer enjoy, making all future wars *come as you are*. As this new policy of military preparedness evolved, so did the role of the operations analyst. Today's military establishment has accepted the idea that analytic techniques can be applied to military policy, strategy and tactics on a regular basis.

1.1 Background

The modern operations analyst assists decision makers in allocating resources for the planning, development, acquisition and use of military systems. In the current environment of radical force reductions, effective and efficient allocation is critical. Rather than recommend a specific course of action, the analyst best serves the decision maker by providing insight into competing alternatives. This is often accomplished through the analyst's ability to quantify complex interrelationships and reduce unmanageable amounts of information. Factor analysis is an analysis tool well suited to this type of task. Originally developed to study behavioral and biological phenomenon, factor analysis is a statistical technique for simplifying complex relationships that exist among a set of observed variables. It accomplishes this by identifying latent common factors within a data set. By linking seemingly unrelated variables, these factors provide insight into the interrelationship and structure of the data (Dillon and Goldstein, 1972:53).

1.2 Problem Statement

Although factor analysis, the technique, goes back to the turn of the century, (Jackson, 1991:388) it is the increased computational capability of modern computer systems that has its application practical. While this made factor analysis very accessible, it also led to the proliferation of many different factor analysis techniques and variations. Subjective selection of a particular technique can lead to results strikingly different from that of another researcher investigating the same data. Because of this failure to render a unique, reproducible solution, many researchers consider factor analysis unsatisfactory as a scientific methodology.

1.3 Research Objectives

1.3.1 Objective. The primary objective of this research effort was to develop and use a Monte Carlo simulation to assess the sensitivity of factor analysis accuracy in response to such effects as sample size, variable to factor ratio, factor loading strengths and structure complexity. The simulation was written in FORTRAN and allows for any modifications necessary to conduct future sensitivity analysis beyond the scope of this effort.

1.3.2 Secondary Objectives. Methods to measure and characterize the behavior of data to be factor analyzed were developed. Quantifying such behavior provides insight into the accuracy to expect from different factor analysis techniques. This information can provide guidelines for methodology selection, and limit the amount of subjectivity introduced by the analyst conducting factor analysis activities.

II. Literature Review

Factor analysis saw most of its early application in psychology where it became a valuable tool in measuring underlying differences in aptitude as measured by individual test scores (Frutcher, 1959:vii). Most information concerning the evolution of factor analysis is therefore concentrated in literature specific to this area. In addition to its inherent subjectivity, perhaps it is also the association with this science that causes many to view factor analysis with suspicion. Edward Jackson summarizes this very well in his User's Guide to Principal Components when he says,

> ...the solutions were not unique and the early reputation of factor analysis was felt by many in the statistical community to be closely associated with witchcraft" (Jackson, 1991:388).

2.1 Principal Components Analysis

Similar in nature to factor analysis, principal components analysis is another statistical method that can be applied to a group of variables. They are alike in that both analyze the inter correlation among variables within a single set rather than external criteria as in other multivariate techniques. Both attempt to reduce the number of variables into subgroups. However, principal components analysis performs this reduction by deriving a small set of linear combinations of the original variables that explains as much of the total variance of the data as possible. Conversely, factor analysis concentrates on the variance shared by the variables. Simply put, principal components analysis views unobservable factors as a function of the variables, while factor analysis treats the variables as a function of the unobservable factors (Bauer, 1992).

2.2 Common Factor Analytic Approach

The factor analysis model assumes a set of intercorrelated variables with each variable composed of both common and unique parts. The common part, called communality, is the portion of a variable's variance that is shared with other variables in the data set. The unique part, known as uniqueness, is that variance belonging to this variable alone. Any variance due to measurement or experimental error would also contribute to the unique part of the variable. Shared variance provides a means of forming the data into coherent relatively independent subgroups. By studying the structure of the subgroups we can gain insight into the process that generated the combination of outcomes measured by the variables. In addition to observing the relationship of the variables, factor analysis also enables us to reduce a large data set to a smaller, more manageable set of unobservable variables known as common factors (Tabachnick and Fidell, 1983:374). The strength of the relationship (correlations) between the observable variables and unobservable factors is indicated by values known as factor loadings. An algebraic representation of a two factor model for a four variable data set is given below.

$$X_1 = \lambda_{11} f_1 + \lambda_{12} f_2 + e_1 \tag{2.1}$$

$$X_2 = \lambda_{21} f_1 + \lambda_{22} f_2 + e_2 \tag{2.2}$$

$$X_3 = \lambda_{31} f_1 + \lambda_{32} f_2 + e_3 \tag{2.3}$$

$$X_4 = \lambda_{41} f_1 + \lambda_{42} f_2 + e_4 \tag{2.4}$$

where

λ_{ij} is the correlation of variable X_i with factor f_j,

$i = 1, ..., p$

and

$j = 1, ..., q$.

Also, the sum squared factor loadings for X_i

$$h_i^2 = \sum_j \lambda_{ij}^2 \tag{2.5}$$

is the variable's communality, the variance accounted for by all the factors. Therefore,

$$\Phi_i = 1 - h_i^2 \tag{2.6}$$

is the uniqueness, the variance not accounted for by the factors. The common factor analytic model is normally summarized as follows:

$$X = \Lambda f + \Phi \tag{2.7}$$

where

X is the $p \times 1$ vector of observed responses,
Λ is the $p \times q$ matrix of unknown constants (factor loadings),
f is the $q \times 1$ vector of unobservable variables (common factors),

and

Φ is the $p \times 1$ vector of unobservable variables (unique factors).

The process of calculating the correlations (factor loadings) between the observable variables and the unobservable factors is known as factor extraction. Because the calculations for factor extraction are involved and complex, this section is limited to the fundamental equations necessary to show the general approach for extracting factors from a data set. The procedure is outlined below:

- First, the data set should consist of several subjects each of which has been measured on p variables.
- Generate Σ, the $p \times p$ data correlation matrix
- Calculate the associated eigenvectors and eigenvalues

$$\Sigma = VLV^T \tag{2.8}$$

where

Σ is the $p \times p$ correlation matrix,

V is the $p \times p$ matrix of eigenvectors

and

L is the $p \times p$ diagonal matrix of eigenvalues.

- Finally, extract the $p \times p$ factor loading matrix

$$\Lambda = V\sqrt{L} \tag{2.9}$$

where

V is $p \times d$, d being the number of retained factors (dimensionality).

The reader should note that the eigenvalues describe the total variance explained by the respective common factors. Eigenvalues can be calculated from a factor loading matrix by summing the square of the factor loadings for each column as

$$\sum_{j} \lambda_{ij}^{2}. \tag{2.10}$$

2.3 Dimensionality Assessment

Prior to extracting the factor loading matrix, the analyst must make a dimensionality assessment in terms of deciding how many of the f_j factors should be retained. The number of eigenvalues, eigenvectors and corresponding common factors will equal the number of variables in the data set. Therefore, while reducing the dimensionality of the data set is desirable, it must be done without jeopardizing our ability to interpret the data structure. Several methods have been put forth to assist in dimensionality assessment, including the scree test (Catell, 1966), Horn's test (Horn, 1965) and Kaiser's criteria (Kaiser, 1960). Kaiser's criteria is one of the simplest methods in that it suggests discarding all factors with an eigenvalue less than one. For detailed information on dimensionality assessment, especially as it applies to factor analysis, see (Bauer, 1981).

2.4 Exploratory versus Confirmatory Factor analysis

Factor analysis is normally conducted in one of two contexts. In the exploratory context, the researcher has no idea what the underlying structure of the data may be and conducts factor analysis for the purpose of gaining insight into that structure. If the researcher has no prior information concerning the underlying structure of the data or is testing some hypothesis for the structure, he is conducting confirmatory factor analysis.

2.5 Factor Interpretation

When interpreting a factor pattern, we investigate the λ_{ij} loadings of the factor pattern matrix so as to identify factor-variable correlations and to determine the strength of these correlations. A commonly used interpretation heuristic applied against an example two factor by six variable model (Lawley and Maxwell, 1971) in Table 2.1 demonstrates factor interpretation.

- For each variable, move across the factors (row) selecting the loading value with the largest absolute value
- Assess each selected loading for significance. A common rule of thumb requires a loading to be greater than .30
- Address each variable not loaded on a factor and decide if, based on its importance to the analysis, the variable should be dropped
- Assign some intuitive meaning to the pattern of selected loadings that describes the meaning of the variables that load on each of the factors

Table 2.1. Example Factor Pattern Interpretation

Variables	f_1	f_2	h_i^2	Φ_i^2
Gaelic	**.553**	.429	.490	.510
English	**.568**	.288	.406	.594
History	.392	**.450**	.356	.644
Arithmetic	**.740**	.-273	.623	.377
Algebra	**.724**	.-211	.569	.431
Geometry	**.595**	.-132	.372	.628

If this factor pattern had been extracted from data corresponding to six different sets of test scores, for example, the researcher might infer that History is influenced by a mental process independent from that influencing the other subjects.

2.6 *Simple Structure*

Unfortunately, not all loading matrices possess structures so simple as to allow this kind of interpretation procedure. A variable often has moderate loadings on more than one factor. In these situations, pattern interpretation can become very subjective and difficult to reproduce. If we could reduce the number of significant loadings for a given variable, hopefully the structure would be simplified and easier to interpret. Before addressing the various methods for simplifying a factor pattern, it is necessary to decide upon some type of criterion for simple structure. One of the first and still widely accepted criterion was suggested by Thurstone in 1932 (Kaiser, 1958:187). This criterion for a $p \times q$ loading matrix consists of the following five main points (Morrison, 1990:368):

- Each row should contain at least one zero
- Each column should contain at least q zeros
- Every pair of columns should contain several responses whose loadings vanish in one column but not in the other
- If the number of factors is four or more, every pair of columns should contain a large number of responses with zero loadings in both columns
- Conversely, for every pair of columns only a small number of responses should have nonzero loadings in both columns

Since simple structures meeting all of Thurstone's conditions are rarely found in real world data, most simplification techniques emphasize different aspects of Thurstone's criterion. These techniques fall into a class of methodologies known as factor rotations.

2.7 *Factor Rotation*

A factor rotation is any transformation of eigenvectors or original variables that seeks to simplify pattern structure (Jackson, 1991:167). Carroll proposed one of the first analytic criterion for approximating simple structure. His method minimized the sums of cross products (across factors) of squares of factor loadings (Carroll, 1953). While arriving at a unique solution and approximating a simple structure, Carroll's method was very computationally complex. Howe, and Lawley and Maxwell proposed least squares methods for rotating as closely as possible to a pattern of zero and nonzero loadings. Joreskog developed a methodology utilizing hypothesis testing for simple structure (Morrison, 1990:369). A factor pattern and its associated rotated factor pattern are shown below in Table 2.2 to illustrate the structure simplification that can be accomplished through factor rotation. Although they have identical communality and uniqueness, note how using our simple heuristic, the rotated pattern on the right is far easier to interpret than the original pattern on the left.

Table 2.2. Example Unrotated versus Rotated Factor Pattern

Variables	f_1	f_2	h_i^2	Φ_i^2	f_1	f_2	h_i^2	Φ_i^2
Gaelic	**.55**	.42	.49	.51	.36	**.59**	.49	.51
English	**.56**	.28	.40	.59	.43	**.46**	.40	.59
History	.39	**.45**	.35	.64	.21	**.55**	.35	.64
Arithmetic	**.74**	-.27	.62	37	**.78**	.00	.62	.37
Algebra	**.72**	-.21	.56	.43	**75**	.05	.56	.43
Geometry	**.59**	-.13	.37	.62	**.60**	.08	.37	.62

Also note that the structure of the rotated pattern has changed from that of the original factor pattern. A researcher might now infer that two separate processes underlie the test scores. Namely, language and History are creative or right brain activities, while the remaining subjects are analytical or left brain activities. As mentioned, simplification can reduce some of the subjectivity in structure interpretation. But we also said each rotation technique may choose to accomplish simple structure through a different aspect of Thurstone's criterion. Therefore, what was intended to introduce objective, scientific methodology into the process has aggravated the problem and generated controversy. David Saunders wrote:

> While some have argued that computers should be used to do factor analysis better, rather merely more quickly, it has been difficult to define what better should mean. Rotation is one of the difficult cases in point. Since computers have become available, there have been no less than fifteen distinguishable approaches to formulations of simple structure that are more or less amenable to computer programming. (Saunders, 1960:199)

The thirty years since this statement have seen little improvement. In his Ph.D. thesis, Henry F. Kaiser addressed some of the original attempts to strip away the subjectivity of factor rotation so factor analysis could become a proper tool for scientific inquiry. All of these attempts employed some form of mathematical condition for simple simple structure.

2.8 Quartimax Criterion

Neuhaus and Wrigley, Saunders and Ferguson simultaneously suggested one of the first rotation criterion for simple structure (Kaiser, 1958:188). Known as the Quartimax criterion, it seeks to maximize the sum of the fourth powers of the rotated loadings. This is the mathematical equivalent of attempting to place only one major loading on a given factor. The quartimax criterion remains one of the most popular and accessible methodologies today. Its formulated as

$$max \frac{\left[pq \sum_{i=1}^{p} \sum_{j=1}^{q} (\lambda_{ij}^{2})^{2} - (\sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_{ij}^{2})\right]}{p^{2}q^{2}}. \tag{2.11}$$

This was intended to fulfill Thurstone's requirement that a large loading on one factor should be opposite a small loading on any adjacent factor.

2.9 Varimax Criterion

Henry Kaiser criticized the quartimax approach because he felt it concentrated too much on Thurstone's mathematically unmanageable, qualitative conditions. Proposing a departure from Thurstone's criterion, Kaiser pursued simple structure by maximizing the variance of the loadings across variables within factors. With this approach, loadings tend to become higher for those variables with high correlations with a factor and smaller for the other variables (Tabachnick and Fidell ,1983:398). Kaiser named this criterion the varimax criterion. Varimax is also a very popular rotation procedure and is found in all statistics software packages. Its formulation is

$$max \sum_{j=1}^{q} \left[\sum_{i=1}^{p} \lambda_{ij}^{4} - \frac{(\sum_{i=1}^{p} \lambda_{ij}^{2})}{p} \right]. \qquad (2.12)$$

2.10 Normal Varimax Criterion

Kaiser conducted numerous experiments to assess the properties of both the quartimax and varimax criterion. In these experiments, he generated rotated factor patterns with the quartimax and varimax criterion from data produced by a series of twenty-four psychological tests administered by Holzinger and Harman (Kaiser, 1958:192). Given to a set of forty-five school children, these tests were designed to measure such aptitudes as visual perception, paragraph comprehension, word comprehension, and problem reasoning (Dillon and Goldstein, 1984:310). The correlations and structure of this data were well documented, giving Kaiser *a priori* knowledge as to the structure each rotation criterion needed to recover. Kaiser found that quartimax tended to create a large general factor, but varimax recovered the Holzinger and Harman factor pattern fairly well. However, he noted a bias in the varimax loadings in which all loadings of the factors with the larger variance contributions were larger than the loadings for the smaller factors (Kaiser, 1958:192). Recall from the previous section on the factor analytic approach that variance contribution of a factor is explained by that factor's associated eigenvalue. Kaiser attributed the observed bias to the fact that the varimax criterion was a fourth-power function. He reasoned that each variable in the unrotated factor pattern is weighted by its communality, and that in a fourth-power function, a variable with a communality of 0.6, for example, would be influenced by the rotation criterion four times as much as a variable with a communality of 0.3 (Kaiser, 1958:193). As discussed, the communality of a variable is the sum squared factor loadings for that variable. Assuming a one factor model, a variable with a communality of 0.6 would have a corresponding factor loading of 0.77. Similarly the communality 0.3 has a corresponding factor loading of 0.54. When we take the fourth power of the factor loadings, the value resulting from the larger loading is four times that of the smaller loading, or .35 versus .085.

To compensate for this bias, Kaiser normalized the extracted factors to reduce the relative influence of each test during the rotation. This consisted of dividing each variable's loading by the square root of its communality. After rotation, these weighting factors are removed (Kaiser, 1958:193). From this point Kaiser called the weighted version of the varimax criterion a normal varimax, and the original version subsequently became known as raw varimax. The formulation for normal varimax is

$$max \sum_{j=1}^{q} \left[\sum_{i=1}^{p} (\lambda_{ij}^{*})^4 - \frac{(\sum_{i=1}^{p} (\lambda_{ij}^{*})^2)}{p} \right] \tag{2.13}$$

where

λ_{ij}^{*} is equal to λ_{ij} weighted by the communality of variable X_i,

$$\lambda_{ij}^* = \frac{\lambda_{ij}}{h_i^2}. \tag{2.14}$$

Kaiser confessed to having no rational basis for choosing this weighting scheme, and admitted making a numerical-intuitive selection. Indeed, the normal varimax criterion did recover the Holzinger and Harman structure without the disturbing bias found in the raw varimax solutions. One side effect of the normal varimax criterion, however, was that it reduced the divergence of the eigenvalues. We contend that significant alteration of eigenvalues removes important information regarding the significance of the factors and thus the variance explained by the factors. For this reason, eigenvalue recovery for the Holzinger and Harman data should have been included as a measure of effectiveness when the normal varimax criterion was evaluated.

2.11 Monte Carlo Studies

The properties of many models and computational algorithms have become fairly well understood as a result of Monte Carlo studies. A Monte Carlo study is a class of study that employs random variates to solve statistical problems that are insensitive to the passage of time. These statistical problems are usually not analytically tractable (Law and Kelton, 1991:113-114).

2.11.1 MacCallum Study. In 1977 MacCallum conducted a Monte Carlo study to determine the the ability of ALSCAL, a computerized algorithm, to recover true structure inherent in simulated proximity measures (MacCallum, 1977). ALSCAL attempts to identify separate homogeneous subgroups of individuals within a sample, then represent each subgroup by a real or hypothetical individual. The solution is intended to provide information about stimulus relationships characterizing different subgroups. In their study they generated, for a sample of individuals, a set of simulated proximity values with a known structure. After analyzing these values with the ALSCAL procedure they measured the degree to which ALSCAL was able to recover the inherent structure. The number of stimuli, the number of individuals, the dimensionality, and the level of random error were all varied. In addition to structure, they also measured ALSCAL's ability to recover stimulus dimension and individual stimulus weights. One interesting discovery was that over the range of their study, the number of individuals had no influence on the other measures of recovery.

2.11.2 Cliff and Pennell Study. In this study Cliff and Pennell used a Monte Carlo approach to determine if certain variables produced systematic effects on the sampling variability of individual factor loadings (Cliff and Pennell, 1967). They believed a researcher's greatest concern to be whether or not to expect a radically different structure for the same variables when obtaining another sample from the same population and whether a given loading was statistically significant. In this effort they contrived two experimental factor matrices and multiplied each by its transpose to produce a correlation matrix. The contrived factor matrices were created with four different factor strengths, four different communalities, and four different loading sizes. In turn, each correlation matrix was used to produce fifty sample factor matrices which were then rotated ,using a least squares fit, to the population values. After rotation, Cliff and Pennell measured the standard deviation of the sample loadings to determine their variability in light of the independent parameters. They observed no consistent relation between sample size or loading size and factor loading stability. Cliff and Pennell did discover that greater communalities led to greater factor loading stability.

2.11.3 Pennell Study. Pennell again undertook a Monte Carlo approach in 1968 to further explore the dependence of factor loading sampling error on sample size and the communality of the variable on which the loading occurs (Pennell, 1968).

In this study, Pennell calculated the sample factor matrices using a method slightly different from that of his work with Cliff . Instead of contriving the experimental factor matrices, Pennell generated them randomly. He then created variables with different communalities and inserted them into randomly generated factor matrices. His work found that increasing communality or sample size resulted in decreased variability in individual loadings. For near zero loadings, sample size had the largest influence.

2.11.4 Tucker, Koopman and Linn Study. This study used a Monte Carlo procedure to study the effectiveness of factor analytic methods using simulated correlation matrices that were more similar to real data correlation matrices than those computed directly from factor patterns. In this approach, Tucker started with a factor pattern that he called the formal model. He then confounded his formal model with the addition of numerous random influence factors and error terms to arrive at a pattern he considered closer to real world. This formal model was then multiplied by its transpose to yield a correlation matrix which was then factor analyzed to yield a factor loading matrix. This matrix was known as the simulation model. The simulated correlation matrices were designed to exhibit different numbers of factors (dimensionality) and different proportions of variance (eigenvalues) explained by the factors. When the simulation model was compared to the formal model, large differences in the quality of results were noted with fewer factors and larger proportions of factor variance.

2.11.5 Bauer Study. Bauer's Monte Carlo study was initiated to characterize the properties of different dimensionality techniques (Bauer, 1981). He addresses the bias effect of sample size, number of variables, number of factors, and structure complexity on the Kaiser and Catell Scree tests for dimensionality. Bauer also generates sample factor loading matrices by factor analyzing and rotating correlation matrices computed from the product of a contrived factor loading matrix and its transpose. He determines the effectiveness of each dimensionality procedure by calculating the mean square error between every contrived factor loading pattern and its associated simulated matrix. His work demonstrated that mean square error could be predicted with fair accuracy.

2.11.6 Summary. This chapter examined the development and evolution of factor analysis as a multivariate analysis tool. This included discussion of the common factor analytic model and several rotation techniques for simplifying structure. Particular attention was paid to the problems of subjectivity and proliferation of factor rotation criteria. We also reviewed many efforts to understand the properties of different factor analysis techniques and to identify influential effects. A Monte Carlo approach was central to nearly all of these studies. The reader should note that there was no evidence of this type of study being conducted to compare and evaluate rotation criterion. True, Kaiser and others did some limited comparison of the structures returned by the quartimax, raw varimax and normal varimax criterion, but these were extremely narrow in scope. Recall that Kaiser admits to no rational basis for selecting his method of normalizing factor loadings. In addition, he compares his new criterion to raw varimax against one type of factor pattern. Remember also that his normal criterion tends to pull the eigenvalues closer together, diluting information concerning relative factor strengths and significance. Overall, Kaiser's work raises some interesting questions:

- How well does the normal varimax perform over a wide range of sample size, structure complexity, variable communality and factor loading?
- In addition to recovering the known structure of an experimental factor pattern, how effective is normal varimax in terms of returning variable loadings and eigenvalues?
- With what types of data will normal varimax not outperform raw varimax?
- Are there circumstances where neither rotation criterion is desirable?

These questions will be addressed throughout the course of this effort.

III. Methodology

Chapter three presents the methodology for assessing the accuracy and associated properties of the raw and normal criterion for varimax rotation. This methodology is designed to investigate the impact of sample size, uniqueness and structure complexity on accuracy. The first section covers experimental design and rationale, followed by a description of the data generation techniques. Section 3 proposes predictors and measures of effectiveness. The chapter concludes with a discussion of the analysis conducted to explore the relationship between the suggested measures of effectiveness and behavior descriptors.

3.1 Experimental Design

Recall from chapter 2 that Kaiser based his normal varimax rotation criterion on a weighting scheme that enabled him to restore the structure underlying Holtzinger and Harman's data from twenty-four psychological tests. It must be reemphasized that his validation efforts for this weighted varimax were limited to that structure only. Table 3.1 includes the factor loading matrix for Holzinger and Harmon's tests and an additional matrix designed to show the pattern structure.

This matrix was generated by Holtzinger and Harman through subjective rotation criterion and is accepted as the true structure for the tests. The twenty-four psychological tests were combined to measure five separate aptitudes among grade school children. Five underlying factors are clearly evident in the structure pattern matrix. Kaiser considered this, or any structure with the variables loaded roughly equal on the factors, a complex structure. Alternately, he suggests that a structure with the all variables loaded on one factor to be the most simple possible structure. This would agree with Thurstone's conditions for simple structure outlined in the previous chapter. In light of the information concerning Kaiser's normal varimax development and validation, we thought it prudent to investigate the sensitivity of

Table 3.1. Factor Loadings And Pattern Structure For Holtzinger and Harman's Twenty-Four Psychological Tests

A	f_1	f_2	f_3	f_4
X_1	.10	.32	**.62**	.20
X_2	.07	.15	**.41**	.13
X_3	.10	.12	**.53**	.13
X_4	.15	.18	**.53**	.12
X_5	**.75**	.15	.26	.15
X_6	**.72**	.05	.28	.25
X_7	**.81**	.08	.27	.11
X_8	**.54**	.26	.38	.14
X_9	**.76**	.-04	.29	.30
X_{10}	.28	**.66**	.-19	.14
X_{11}	.27	**.61**	.-04	.29
X_{12}	.13	**.72**	.09	.03
X_{13}	.24	**.63**	.31	.02
X_{14}	.23	.19	.-02	**.48**
X_{15}	.11	.14	.08	**.50**
X_{16}	.05	.22	.34	**.45**
X_{17}	.15	.24	.-03	**.62**
X_{18}	.01	.39	.20	**.52**
X_{19}	.12	.22	.18	**.39**
X_{20}	.31	.18	**.46**	.29
X_{21}	.17	**.46**	.33	.24
X_{22}	.31	.12	**.40**	.40
X_{23}	.31	.29	**.54**	.25
X_{24}	.39	**.46**	.14	.31

A^*	f_1	f_2	f_3	f_4
X_1			X	
X_2			X	
X_3			X	
X_4			X	
X_5	X			
X_6	X			
X_7	X			
X_8	X			
X_9	X			
X_{10}		X		
X_{11}		X		
X_{12}		X		
X_{13}		X		
X_{14}				X
X_{15}				X
X_{16}				X
X_{17}				X
X_{18}				X
X_{19}				X
X_{20}			X	
X_{21}		X		
X_{22}			X	
X_{23}			X	
X_{24}		X		

Kaiser's criterion against structures of varying complexity. As seen in Tables 3.2, 3.3 and 3.4, we began the research effort by constructing nine initial 16 variable by 4 factor experimental factor patterns.

Observe the three different levels of structure complexity for these experimental factor patterns:

- Level 1: Variables loaded equally on all four factors
- Level 2: Variables predominantly loaded on the first two factors

Table 3.2. Initial Level 1 Theoretical Factor Patterns

Pattern 1			
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
0	.9	0	0
0	.9	0	0
0	.9	0	0
0	.9	0	0
0	0	.9	0
0	0	.9	0
0	0	.9	0
0	0	.9	0
0	0	0	.9
0	0	0	.9
0	0	0	.9
0	0	0	.9

$h_i^2 = .81$

Pattern 2			
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	0	.8	0
0	0	.8	0
0	0	.8	0
0	0	.8	0
0	0	0	.8
0	0	0	.8
0	0	0	.8
0	0	0	.8

$h_i^2 = .64$

Pattern 3			
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	0	.7	0
0	0	.7	0
0	0	.7	0
0	0	.7	0
0	0	0	.7
0	0	0	.7
0	0	0	.7
0	0	0	.7

$h_i^2 = .49$

Table 3.3. Initial Level 2 Theoretical Factor Patterns

Pattern 4			
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
0	.9	0	0
0	.9	0	0
0	.9	0	0
0	.9	0	0
0	.9	0	0
0	0	.9	0
0	0	.9	0
0	0	0	.9
0	0	0	.9

$h_i^2 = .81$

Pattern 5			
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	0	.8	0
0	0	.8	0
0	0	0	.8
0	0	0	.8

$h_i^2 = .64$

Pattern 6			
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	0	.7	0
0	0	.7	0
0	0	.7	0
0	0	.7	0

$h_i^2 = .49$

Table 3.4. Initial Level 3 Theoretical Factor Patterns

Pattern 7					Pattern 8					Pattern 9				
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
.9	0	0	0		.8	0	0	0		.7	0	0	0	
0	.9	0	0		0	.8	0	0		0	.7	0	0	
0	.9	0	0		0	.8	0	0		0	.7	0	0	
0	0	.9	0		0	0	.8	0		0	0	.7	0	
0	0	.9	0		0	0	.8	0		0	0	.7	0	
0	0	0	.9	$h_i^2 = .81$	0	0	0	.8	$h_i^2 = .64$	0	0	0	.7	$h_i^2 = .49$

- Level 3: Variables predominantly loaded on the first factor

The eigenvalues indicate the total variance explained by each factor. One can readily see that our different levels of structure complexity would affect the divergence of the eigenvalues, with the simplest structure producing the largest divergence. The three levels of structure complexity were chosen to investigate the impact of structure complexity as manifested in eigenvalue divergence. In order to quantify Kaisers's subjective description of structure complexity, we proposed using the ratio of highest to lowest eigenvalue of each experimental factor pattern. We will refer to this value as the *stability ratio*. Since eigenvalues provide important information on the validity and strength of a factor, varying stability ratios will hopefully provide insight as to each criterion's ability to restore this information upon factor pattern rotation. It was noted in chapter 2 that Kaiser did not incorporate eigenvalue recovery as a measure of effectiveness for normal varimax rotation. For that matter, varimax seemed to have a leveling effect on eigenvalue divergence. Our range of structure complexity provides a range of eigenvalue dispersion for assessing rotational properties. Notice at each level of structure complexity, the initial nine experimental factor patterns exhibit three different average communalities of .81, .64, and .49. This will provide insight into the influence of structure complexity, communality and the interaction of both.

As shown in Tables 3.5 thru 3.10, each of the initial nine experimental patterns were then perturbed across both the factors and the variables.

Table 3.5. Level 1 Perturbed Across Factors

Pattern 10			
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	0	.7	0
0	0	.7	0
0	0	.7	0
0	0	.7	0
0	0	0	.6
0	0	0	.6
0	0	0	.6
0	0	0	.6

Pattern 11			
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	0	.6	0
0	0	.6	0
0	0	.6	0
0	0	.6	0
0	0	0	.5
0	0	0	.5
0	0	0	.5
0	0	0	.5

Pattern 12			
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
0	.6	0	0
0	.6	0	0
0	.6	0	0
0	.6	0	0
0	0	.5	0
0	0	.5	0
0	0	.5	0
0	0	.5	0
0	0	0	.4
0	0	0	.4
0	0	0	.4
0	0	0	.4

Note when perturbing across the factors, each variable within a common factor exhibits the same communality. Perturbing across the variables assigns a different communality to each variable within a common factor. In terms of the three subjective levels of structure complexity defined earlier in this section, the structural complexity is maintained to allow the researcher to assess how eigenvalue dispersion due to both variable communality and structure complexity drive rotation accuracy. For the purposes of this effort, rotational accuracy is the ability of a rotation criterion to recover the loadings, eigenvalues and structure pattern of each experimental factor pattern. Measures of effectiveness were suggested to capture these properties

Table 3.6. Level 2 Perturbed Across Factors

Pattern 13			
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	.8	0	0
0	0	.7	0
0	0	.7	0
0	0	0	.6
0	0	0	.6

Pattern 14			
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	.7	0	0
0	0	.6	0
0	0	.6	0
0	0	0	.4
0	0	0	.4

Pattern 15			
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
0	.6	0	0
0	.6	0	0
0	.6	0	0
0	.6	0	0
0	.6	0	0
0	0	.5	0
0	0	.5	0
0	0	0	.4
0	0	0	.4

Table 3.7. Level 3 Perturbed Across Factors

Pattern 16			
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
.9	0	0	0
0	.8	0	0
0	.8	0	0
0	0	.7	0
0	0	.7	0
0	0	0	.6

Pattern 17			
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
.8	0	0	0
0	.7	0	0
0	.7	0	0
0	0	.6	0
0	0	.6	0
0	0	0	.5

Pattern 18			
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
.7	0	0	0
0	.6	0	0
0	.6	0	0
0	0	.5	0
0	0	.5	0
0	0	0	.4

Table 3.8. Level 1 Perturbed Across Factors And Variables

Pattern 19

.9	0	0	0
.8	0	0	0
.7	0	0	0
.6	0	0	0
0	.8	0	0
0	.7	0	0
0	.6	0	0
0	.5	0	0
0	0	.7	0
0	0	.6	0
0	0	.5	0
0	0	.4	0
0	0	0	.6
0	0	0	.5
0	0	0	.4
0	0	0	.3

Pattern 20

.8	0	0	0
.7	0	0	0
.6	0	0	0
.5	0	0	0
0	.7	0	0
0	.6	0	0
0	.5	0	0
0	.4	0	0
0	0	.6	0
0	0	.5	0
0	0	.4	0
0	0	.3	0
0	0	0	.5
0	0	0	.4
0	0	0	.3
0	0	0	.3

Pattern 21

.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
0	.6	0	0
0	.5	0	0
0	.4	0	0
0	.3	0	0
0	0	.5	0
0	0	.4	0
0	0	.3	0
0	0	.3	0
0	0	0	.4
0	0	0	.4
0	0	0	.3
0	0	0	.3

Table 3.9. Level 2 Perturbed Across Factors And Variables

Pattern 22

.9	0	0	0
.8	0	0	0
.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
0	.8	0	0
0	.7	0	0
0	.6	0	0
0	.5	0	0
0	.4	0	0
0	0	.7	0
0	0	.6	0
0	0	0	.6
0	0	0	.5

Pattern 23

.8	0	0	0
.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
.3	0	0	0
0	.7	0	0
0	.6	0	0
0	.5	0	0
0	.4	0	0
0	.3	0	0
0	0	.6	0
0	0	.5	0
0	0	0	.4
0	0	0	.3

Pattern 24

.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
0	.6	0	0
0	.5	0	0
0	.4	0	0
0	.3	0	0
0	.3	0	0
0	0	.5	0
0	0	.4	0
0	0	0	.4
0	0	0	.3

Table 3.10. Level 3 Perturbed Across Factors And Variables

Pattern 25			
.9	0	0	0
.8	0	0	0
.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
0	.8	0	0
0	.7	0	0
0	0	.7	0
0	0	.6	0
0	0	0	.6

Pattern 26			
.8	0	0	0
.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
0	.7	0	0
0	.6	0	0
0	0	.6	0
0	0	.5	0
0	0	0	.4

Pattern 27			
.7	0	0	0
.6	0	0	0
.5	0	0	0
.4	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
.3	0	0	0
0	.6	0	0
0	.5	0	0
0	0	.5	0
0	0	.4	0
0	0	0	.3

and will be defined in a later section. Further, the mean and standard deviation of these measures of effectiveness will be recorded over 100 experiments for each of the thirty-six populations corresponding to the experimental factor patterns.

Finally, the set of experimental patterns was further expanded by randomly adding nuisance factors, ranging between .64 and .01, to variables of the perturbed factor patterns. This has the effect of inflating the communalities of the corresponding variables. These additional experimental factor patterns are shown in Tables 3.11 thru 3.13.

Each theoretical factor pattern was analyzed over sample sizes of 17, 32, 160 and 500. Rather than determine the improvement in factor analysis effectiveness as sample size increased, we elected to explore the minimum threshold needed to facilitate reasonably accurate factor analysis. The flow diagram in figure 3.1 illustrates a general overview of the approach utilized in the analysis of each theoretical factor pattern. The following section will address each step of this approach in more detail.

Table 3.11. Level 1 Perturbed And Nuisance Loadings

Pattern 28			
.9	0	0	0
.8	0	.2	0
.7	.6	0	0
.6	0	.4	0
0	.8	.3	0
.5	.7	0	0
0	.6	0	.2
.2	.5	0	0
0	0	.7	0
.5	0	.6	0
0	0	.5	0
.2	0	.4	0
0	0	0	.6
0	.4	0	.5
.1	0	0	.4
0	.2	0	.3

Pattern 29			
.8	0	0	.6
.7	0	0	0
.6	.5	0	0
.5	0	0	0
.5	.7	0	0
0	.6	0	0
0	.5	0	.1
0	.4	.3	0
0	0	.6	0
.4	0	.5	0
0	0	.4	0
.2	0	.3	0
0	.4	0	.5
0	0	0	.4
0	.4	0	.3
0	0	.1	.3

Pattern 30			
.7	0	0	0
.6	0	.2	0
.5	0	0	0
.4	0	.1	0
0	.6	0	0
0	.6	0	0
0	.4	0	0
0	.3	.1	0
0	0	.5	0
.2	0	.4	0
0	0	.3	0
.1	0	.3	0
0	0	0	.4
0	.3	0	.4
0	0	0	.3
.1	0	0	.3

Table 3.12. Level 2 Perturbed And Nuisance Loadings

Pattern 31			
.9	0	0	.4
.8	0	0	0
.7	0	.3	0
.6	0	0	0
.5	0	.2	0
.4	0	0	0
.3	.3	0	0
0	.8	0	0
.2	.7	0	0
0	.6	0	0
0	.5	0	.4
0	.4	0	0
.2	0	.7	0
0	0	.6	0
0	.5	0	.6
0	0	0	.5

Pattern 32			
.8	0	0	0
.7	0	.6	0
.6	0	0	0
.5	0	.4	0
.4	0	0	0
.3	.2	0	0
.3	0	0	0
0	.7	0	.4
0	.6	0	0
.4	.5	0	0
0	.4	0	0
0	.3	0	.3
0	0	.6	0
.1	0	.5	0
0	0	0	.4
.2	0	0	.3

Pattern 33			
.7	0	0	.2
.6	0	0	0
.5	.4	0	0
.4	0	0	0
.3	0	.3	0
.3	0	0	0
.3	0	0	.1
0	.6	0	0
.1	.5	0	0
0	.4	0	0
0	.3	.1	0
0	.3	0	0
.4	0	.5	0
0	0	.4	0
0	.2	0	.4
0	0	0	.3

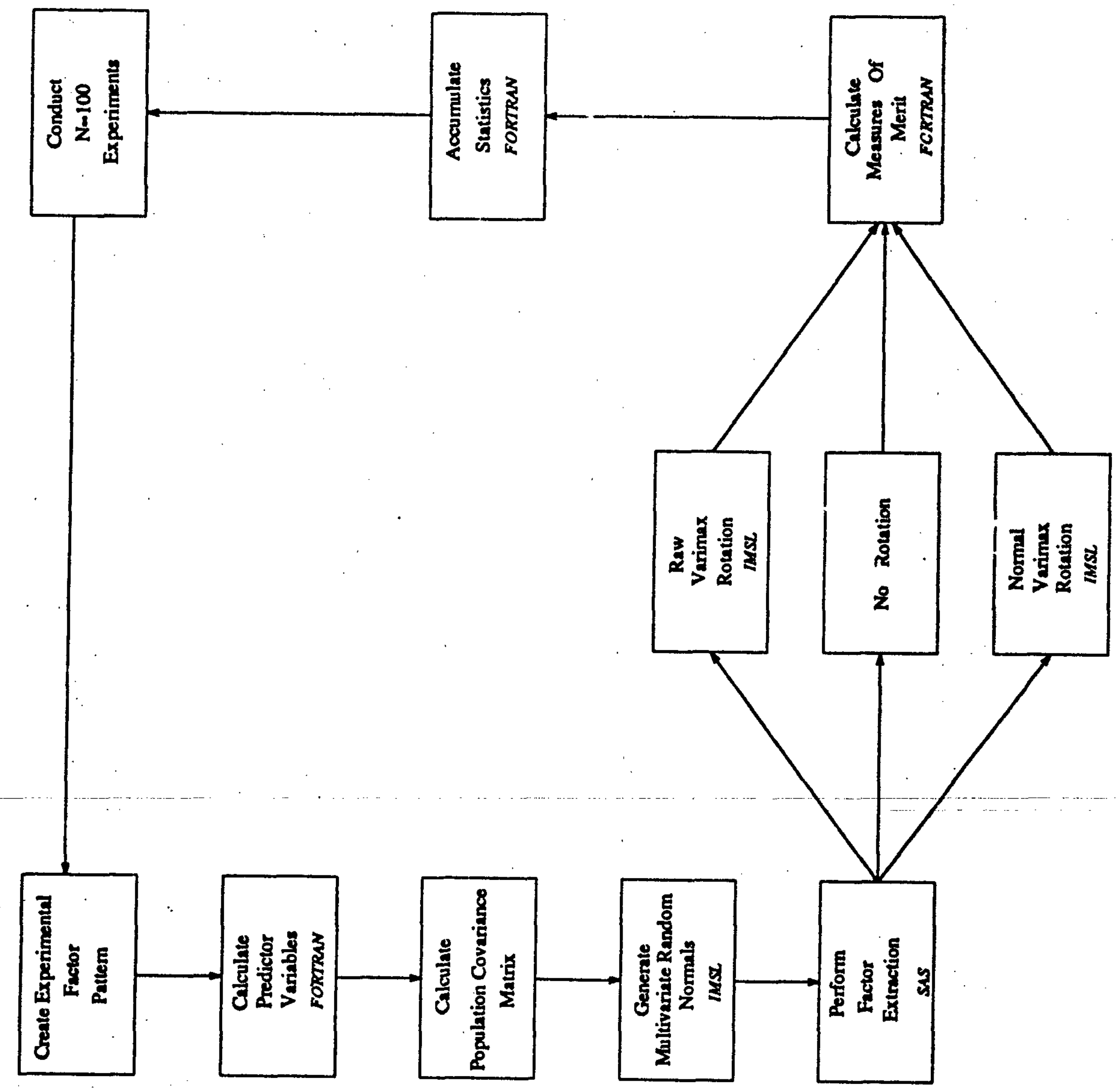


Figure 3.1. Methodology

Table 3.13. Level 2 Perturbed And Nuisance Loadings

Pattern 34			
.9	0	.5	0
.8	0	0	0
.7	.6	0	0
.6	0	0	0
.5	0	.1	0
.4	0	0	0
.3	0	0	.2
.3	0	0	0
.3	0	.3	0
.3	0	0	0
.3	.2	0	0
0	.8	0	0
0	.7	0	.5
0	0	.7	0
.4	0	.6	0
0	0	0	.6

Pattern 35			
.8	0	.2	0
.7	0	0	0
.6	0	.5	0
.5	0	0	0
.4	.2	0	0
.3	0	0	0
.3	0	0	.3
.3	0	0	0
.3	0	.4	0
.3	0	0	0
.3	.2	0	0
0	.7	0	0
0	.6	0	.5
0	0	.6	0
.3	0	.5	0
0	0	0	.4

Pattern 36			
.7	0	.1	0
.6	.5	0	0
.5	0	.2	0
.4	0	0	0
.3	0	.2	0
.3	0	0	0
.3	.1	0	.2
.3	0	0	0
.3	0	.3	0
.3	0	.1	0
.3	0	0	.2
.4	.6	0	0
0	.5	0	0
.1	0	.5	0
0	0	.4	0
0	.2	0	.3

3.2 Data Generation

This section begins by outlining the procedure for deriving a covariance matrix from each of the third-six theoretical factor patterns. Recall the basic factor-analytic model introduced in chapter two,

$$X = \Lambda f + e \tag{3.1}$$

where

X is a $p \times 1$ vector of observed responses,

Λ is a $p \times q$ matrix of unknown constants (factor loadings),

f is a $q \times 1$ vector of unobservable variables (common factors)

and

e is a $p \times 1$ vector of unobservable variables (unique factors).

Specifically,

$$\Lambda = \begin{bmatrix} \lambda_{11} & . & & . & . & . & \lambda_{1q} \\ . & . & & . & . & . & . \\ . & & \lambda_{ij} & & & & . \\ . & & & . & & & . \\ . & & & & & . & . \\ \lambda_{p1} & . & & . & . & . & \lambda_{pq} \end{bmatrix}.$$

Also recall that the j unique factors consist of the variance of each variable not explained by any of the common factors. It is generally assumed that the unique parts of each variable are uncorrelated. Further, the unique parts are uncorrelated with the common parts

$$E = ee^T = \Psi = \begin{bmatrix} \Psi_1 & 0 & . & . & . & 0 \\ 0 & \Psi_2 & . & . & . & 0 \\ . & & . & & & \\ . & & & . & & \\ . & & & & . & \\ 0 & 0 & . & . & . & \Psi_p \end{bmatrix}$$

These assumptions imply that the covariance matrix of X can be written as

$$\Sigma_{xx} = \Lambda \Phi \Lambda^T + \Psi. \tag{3.2}$$

where

Ψ is as previously defined,

Λ is the original experimental factor loadings matrix,

and

$$\Phi = \begin{bmatrix} 1 & & & & \\ \Phi_{21} & 1 & & & \\ \Phi_{31} & \Phi_{32} & 1 & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & & \cdot \\ \Phi_{j1} & \Phi_{j2} & \cdots & \Phi_{q,q-1} & 1 \end{bmatrix}$$

where Φ gives the covariance between the common factors. When it is assumed that the factors are uncorrelated, we have

$$\Phi = I, \tag{3.3}$$

hence

$$\Sigma_{xx} = \Lambda\Lambda^T + \Psi. \tag{3.4}$$

Thus a population covariance matrix can be derived by multiplying the theoretical factor pattern by its transpose and adding the diagonal matrix of uniqueness values for each variable. Since the communality (that part of the variance accounted for by the common factors) is

$$h_i^2 = \sum_j \lambda_{ij}^2, \tag{3.5}$$

The uniqueness of each variable in a given theoretical matrix can be calculated as

$$\Phi_i = 1 - h_i^2 \tag{3.6}$$

Therefore,

$$\Sigma_{xx} = \Lambda\Lambda^T + \left[(1 - \sum_j \lambda_{ij}^2) \right]_{ii}. \tag{3.7}$$

The next step in the experimental procedure was to generate N(sample size) sample vectors of length 16 (number of variables). Since the individual components of each sample vector are obviously not independent,we generated each random vector from a joint (multivariate normal) distribution to form a vector of correlated random normal variates. Each covariance matrix is positive definite and can be factored into an upper and lower triangular matrix using Choleski decomposition

$$\Sigma = C\ C^T \tag{3.8}$$

For a more detailed mathematical explanation of Choleski decomposition, see (Burden and Faires,1989:370). After the covariance matrix was factored, we generated random multivariate normal deviates

$$X_i = \sum_{j=1}^{p} C_{ij} Z_{j1}, \tag{3.9}$$

where

$$i = 1, ..., p$$

and

$$X = (X_1, ... X_p)^T \sim N_p(0, \Sigma). \tag{3.10}$$

A FORTRAN program (see appendix A) utilizing routines CHFAC (IMSL, 1987:1144) and RNMVN (IMSL,1987:1033) was written to read a selected covariance matrix, perform Choleski decomposition and return 100 sets of N(sample size) 1×16 multivariate normal deviate vectors. These data sets were then factor analyzed using the SAS (see appendix C) principal components method. SAS was used because we experienced problems with the FACTR routine (IMSL, 1987:647) returning consistent results. The Monte Carlo technique was implemented in our main FORTRAN program (see appendix D) which used the FROTA routine (IMSL, 1987:656) to return rotated factor patterns for both the raw and normal varimax criterion. There were also problems with the FROTA routine in that the columns of the rotated output were not in the correct order with respect to the eigenvalues. Because of the greater flexibility this routine provided, we decided to use it and created a sort routine in our main program to rewrite the common factors of each rotated factor pattern in the correct (largest eigenvalues first) order. Theoretical factor pattern characterization parameters, rotation measures of effectiveness, and all statistics were calculated in this program.

3.3 Pattern Characterization

Selecting a method that quantifies the fundamental features of a large group of theoretical patterns over a large range of loading structures and communalities was a difficult process. For this effort the researcher relied primarily on three such metrics. They were the

- Stability Ratio
- Complexity index
- Uniqueness index

The stability ratio concept was proposed because it mathematically separated the theoretical factor patterns into three levels of complexity based upon eigenvalue divergence. Patterns where most of the variables load on one factor will display a very large ratio of highest to smallest eigenvalue, thereby incurring a high stability ratio. Conversely, theoretical factor patterns with equal variable loading will have a stability ratio close to 1. Used in this manner, the stability ratio will provide a means of quantifying Kaiser's concept of structure complexity. Since eigenvalue divergence is a function of the factor pattern columns, we will call this component of structure complexity the column effect.

$$\frac{L_{max}}{L_{min}}. \tag{3.11}$$

In addition to eigenvalue divergence, There are several promising methods for quantifying factor pattern complexity. Bauer explains complexity in terms of the difficulty for a factorization technique to extract the first factor over the second (Bauer, 1981:39). In other words, the total variance explained by the first two factors are so close in value, the extraction algorithm has difficulty discerning between the two. He proposed a complexity index combining two structure parts that he called

- Complication due to structure
- Complication due to uniqueness

The first component of the complexity index takes the form

$$\frac{\sum_{i=1}^{p} \sum_{j=2}^{q} \sum_{k=1}^{j-1} (\lambda_{ik}\lambda_{ij})^2}{p} \tag{3.12}$$

This term (structural complexity) suggests that factor patterns with variables loaded equally on all variables will exhibit large complication due to structure with a weak upper bound of 1, while patterns with all variables loaded on one factor will display very low complication due to structure with a lower bound of 0. Inspecting the formulation, we can see that this component is a function of the factor pattern rows and columns. We will call this component of structure complexity the geometry effect. The second part of Bauer's complexity index,

$$\frac{1 - \sum_{i=1}^{p} h_i}{p} \tag{3.13}$$

is the average uniqueness of the experimental factor pattern. Simply put, the lower the uniqueness of a given variable the higher the variance (communality) explained by the common factors. With less variance due to uniqueness (error), a rotation technique has a better chance of recovering the representative factor pattern. We will refer to uniqueness, a function of the rows (variables), as the row effect. In an effort to isolate these two effects, we will perform separated calculations for these two parts. A primary objective of this effort will be to determine if any single effect or interaction of effects can provide accurate prediction of factor pattern behavior.

3.4 Measures Of Effectiveness

Three primary areas were addressed to explore the accuracy of the two varimax criteria.

- The ability to recover the experimental factor pattern structure
- The ability to recover the variable loadings for the experimental factor pattern
- The ability to recover the total variance explained by each factor (eigenvalue) in the experimental factor pattern

Each observed rotated structure was compared to its associated experimental structure to determine the number of times our observed pattern loaded a variable on the same factor as the experimental pattern. This was accomplished by assigning a value of 1 to the most significant factor loading of every variable in each experimental factor pattern and a zero value to all other loadings. This same binary transformation was also performed on each rotated sample factor pattern. The experimental structures were then compared to the M=24 permutations of the j=4 factor columns of its respective rotated solution, scoring the highest number of achieved matches. The scoring calculation is

$$max\ tr\left[\Lambda_e^T\ \Lambda_n\right] \tag{3.14}$$

where

Λ_e is the experimental factor pattern

and

Λ_n (n=1,...,24) are the permuted rotated factor patterns.

The permutations were performed to ensure the study compared only the best possible form of each rotation criterion. Additionally, this same procedure was conducted on each extracted unrotated factor pattern to contrast its accuracy to that of the two varimax rotation criterion

Variable loading recovery was measured by calculating the root mean square residuals between each rotated variable loading and those of its parent experimental factor pattern as

$$\sqrt{\frac{\sum_{j=1}^{q} \sum_{i=1}^{p} (\lambda_{ij}^{*} - \lambda_{ij}^{e})^2}{pq}} \tag{3.15}$$

where

pq is the total number of factor loadings,

λ_{ij}^{e} are the experimental factor pattern loadings

and

λ_{ij}^{*} = are the best permuted factor pattern loadings.

Eigenvalue recovery was measured in a fashion similar to that of the variable loadings. We calculated the root mean square residuals of the eigenvalues between each rotated variable loading and its parent experimental factor pattern. This is calculated as

$$\sqrt{\frac{\sum_{j=1}^{q} (L_j^{*} - L_j^{e})^2}{q}} \tag{3.16}$$

(3.17)

where

q is the total number of eigenvalues,

$\mathbf{L}_j^e$ are the experimental factor pattern eigenvalues

and

$\mathbf{L}_j^*$ are the best permuted factor pattern eigenvalues.

It must be noted that in each of the residual calculations, the rotated factor pattern was the permuted pattern that produced the highest matching score. Again, this was to ensure that residuals were calculated for the best possible rotated representation. All measures of effectiveness were averaged over N=100 data sets from thirty-six different populations corresponding to each experimental factor pattern to establish a grand mean for all measures of effectiveness.

3.5 Regression Analysis

Several regression studies will be conducted to determine if rotated pattern matching score, factor loading error and eigenvalue error can be predicted as functions of sample size, stability, uniqueness and complexity. Each of our measures of merit were taken as the grand mean of 100 iterations for the different experimental structure-sample size combinations. Two types of regression models will be investigated.

1. A linear regression model for predicting measures of effectiveness using a stepwise procedure. Three formulations of this model will include pattern matching score, root mean square residual loadings (RMSR), and root mean square residual eigenvalues (RMSER) as response variables. Each response will be modeled as a linear function of stability ratio, uniqueness index, complexity index and all possible multiplicative interaction terms.

2. A nonlinear model will also be attempted for predicting measures of effectiveness using a stepwise procedure to fit a Cobb-Douglas function. This will enable us to capture any nonlinearities greater than first or second order that may be present in the data. This model will also include three formulations using pattern matching score, RMSR, and RMSER as response variables. The independent variables will include stability ratio, uniqueness index and complexity index as independent variables.

The general approach for the regression analysis will be to first attempt a linear model of each response variable for the smallest sample size, 17. Our criteria for an acceptable model will be any one that explains 80 % of the variance of the response variable as indicated by the adjusted r^2 value. If this model is unacceptable, the sample sizes will be successively increased until the largest, 500, has been modeled. If according to our criterion, the model is still unacceptable, we will attempt to fit a Cobb-Douglas function to the response determined by the smallest sample size. This procedure will also bring in larger samples until all sample sizes are exhausted or an acceptable model is obtained.

3.6 Factor Analysis

Finally, our analysis will conclude with preliminary factor analysis on three data sets consisting of the following groups of variables.

1. The set of variables measuring matching scores for the raw and normal varimax factor patterns, and the unrotated factor pattern.
2. The set of variables measuring RMSR for the raw varimax, normal varimax and unrotated factor patterns.
3. The set of variables measuring RMSER for the raw varimax, normal varimax and unrotated factor patterns.

In addition to generating factor patterns, we will also examine some two-factor plots. This analysis will a visual perspective of the interrelationship of each experimental factor pattern. In order to remain consistent, sample size corresponding to the data set yielding the best regression model for each experimental factor pattern will be the one factor analyzed.

3.7 Summary

This concludes the methodology discussion for this research effort. This chapter covered data generation, the Monte Carlo approach, independent predictor calculations and response measures of effectiveness calculations. The next chapter will detail the results of this approach.

IV. Results and Conclusions

4.1 Monte Carlo Approach

The Monte Carlo method was applied to assess the accuracy and associated properties of the raw and normal varimax criterion for rotation in factor analysis. In all, three different levels of structure complexity were used:

- Level 1: Variables loaded equally on all four factors
- level 2: Variables predominantly loaded on the first two factors
- level 3: Variables predominantly loaded on the first factor

Each level of structure complexity was investigated over a range of average communality and nuisance loadings. As defined in the previous chapter, several factor pattern descriptor variables and measures of effectiveness were proposed to predict rotation criterion performance. Recall that we are defining factor pattern complexity in terms of row, column and geometry effects and that these effects will be estimated as a function of the predictors. Our set of experimental factor patterns were contrived to display a range of each effect. In Figures 4.1, 4.2 and 4.3 we have sorted the set of experimental factor patterns in ascending order of stability ratio, uniqueness index and complexity index to illustrate the experimental range of each effect. Based on the analysis of these metrics, several conclusions have been drawn.

4.2 Pattern Matching

Perhaps one of the most important measures of effectiveness regarding any rotation criterion is the ability to restore the true underlying factor pattern structure. The observed structures for the unrotated and both rotated (raw and normal varimax) factor patterns were compared to each of their associated experimental factor patterns to determine a pattern matching score. That score is the number of times the experimental and observed patterns both load a variable on the same factor.

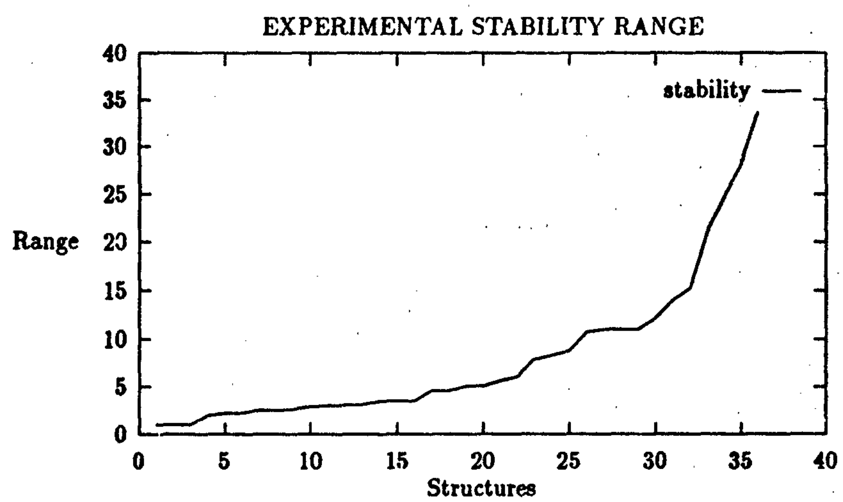


Figure 4.1. Experimental Factor Pattern Stability Ratio Range

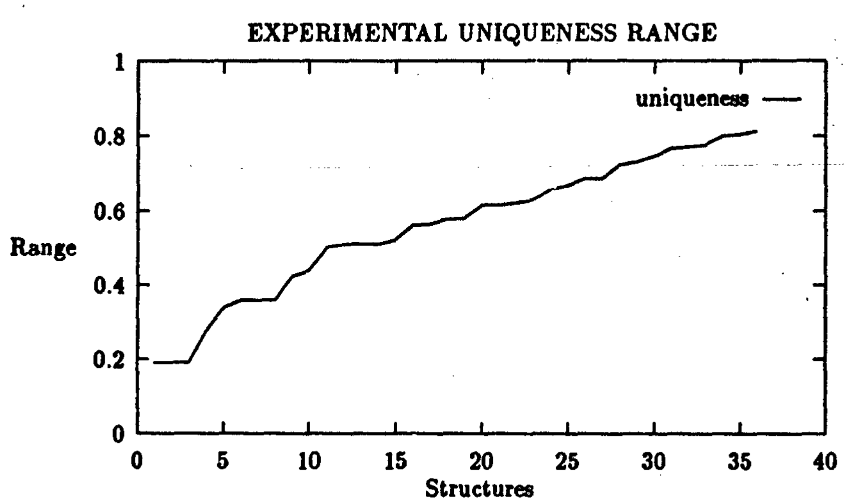


Figure 4.2. Experimental Factor Pattern Uniqueness Index Range

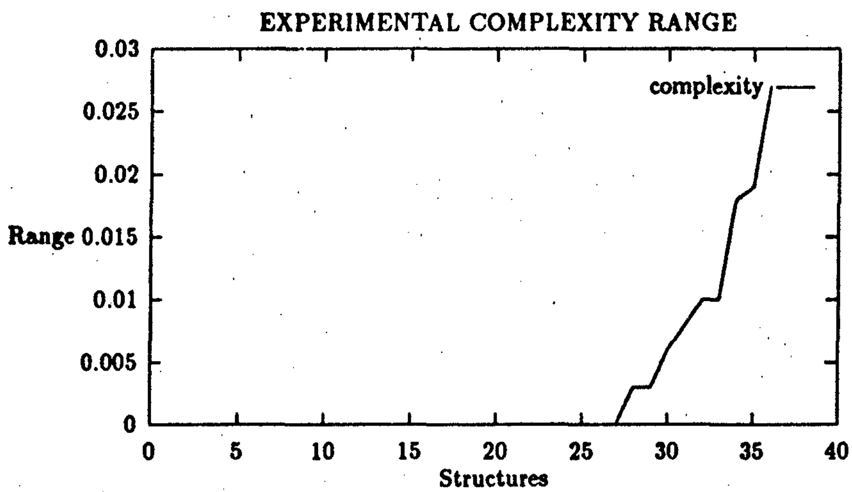


Figure 4.3. Experimental Factor Pattern Complexity Index Range

The pattern matching score was the highest matching total recorded for each of the twenty-four permutations of the four columns in each observed factor pattern. This ensures consistent comparison of the best possible observed factor pattern structure.

4.2.1 Exploratory Scatter Plots. Figures 4.4, 4.5 and 4.6 show initial scatter plots of the matching scores against each complexity effect.

We can clearly see from Figure 4.5 that matching score as a function of uniqueness displays the most conclusive relationship of any predictor variable. Upon closer inspection, their appears to be four different processes taking place. First, there is a general trend toward lower matching scores for the observed unrotated and rotated factor patterns as the average uniqueness of the experimental factor pattern increases. Second, the scatter plot converges with increased uniqueness, indicating no difference between unrotated and rotated matching scores at high levels of uniqueness. Furthermore, we can see that the raw and normal varimax rotation criterion achieve the same matching score at all levels of uniqueness. Finally, Fig-

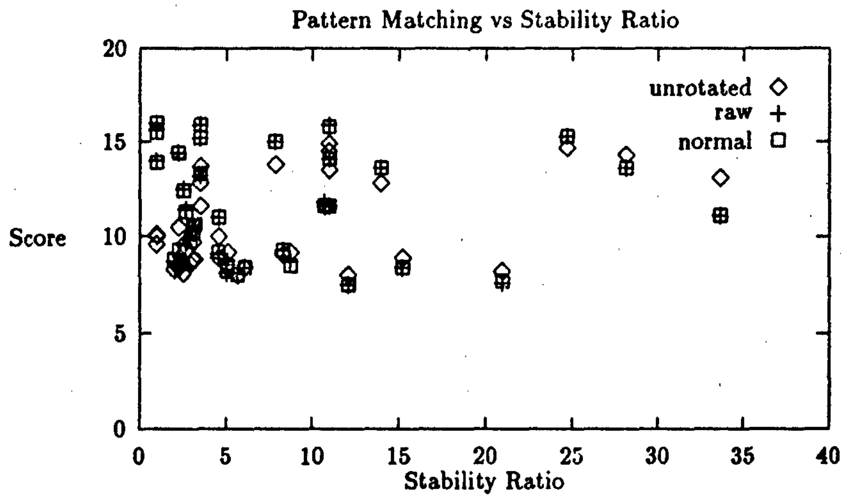


Figure 4.4. Stability Ratio vs Pattern Matching Score For Sample Size N=17

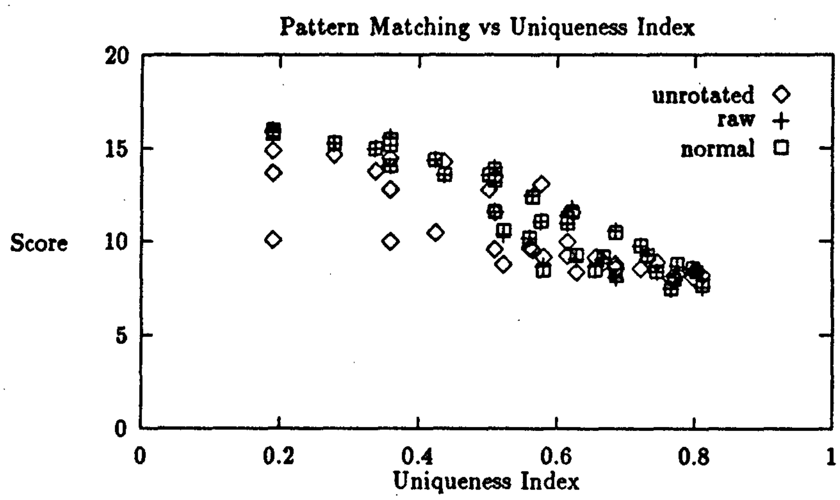


Figure 4.5. Uniqueness Index vs Pattern Matching Score For Sample Size N=17

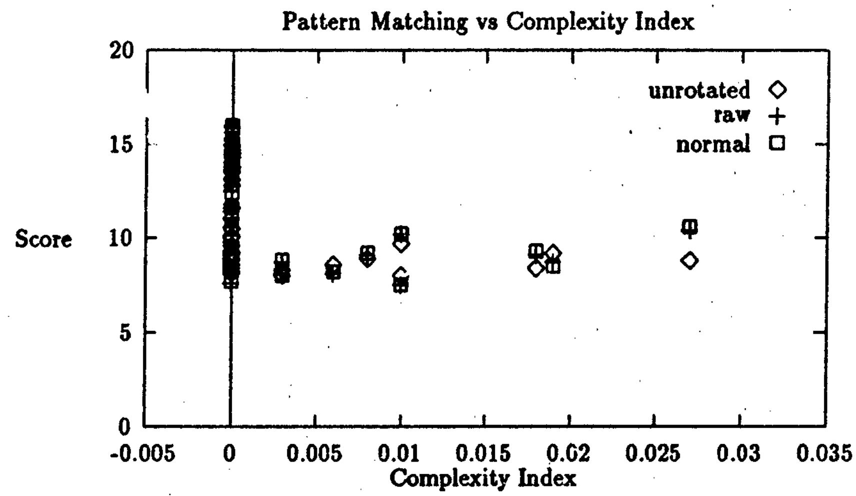


Figure 4.6. Complexity Index vs Pattern Matching Score For Sample Size N=17

ure 4.5 shows unrotated factor patterns of the same uniqueness exhibiting distinctly different matching scores. The significance of this difference decreases as uniqueness increases. To understand what is causing this response, we must refer to the unrotated matching scores ranked according to uniqueness index shown in Table 4.1. This Table shows us that at any given uniqueness, 0.190 for example, higher stability ratios produce a higher matching scores for an unrotated factor pattern. While the scatter plots of stability ratio and complexity index do not show any conclusive relationships, they do indicate that there is no difference between the two varimax criterion in response to column or geometry effects

4.2.2 Regression Analysis. A regression analysis was then performed on the observed matching score for the unrotated and both rotated factor patterns to better understand the influence of column, row and geometry effects as measured by the predictor variables. In addition, a successful regression model will determine if pattern matching scores can be predicted. In this analysis, pattern matching score

Table 4.1. Unrotated Matching Score As A Function Of Uniqueness Index

Structure	Stability	Uniqueness	Score
1	1.000	0.190	10.1
13	3.500	0.190	13.7
25	11.000	0.190	14.9
26	24.750	0.279	14.7
14	7.875	0.339	13.8
5	1.000	0.360	10.0
17	3.500	0.360	12.8
29	11.000	0.360	14.5
2	2.250	0.425	10.5
30	28.160	0.438	14.3
18	14.000	0.502	12.8
9	1.000	0.510	9.6
21	3.500	0.510	11.6
33	11.000	0.510	13.5
4	3.211	0.523	8.8
16	3.097	0.561	9.7
6	2.560	0.565	9.6
34	33.687	0.577	13.1
28	5.108	0.581	9.2
3	2.674	0.615	9.3
15	4.590	0.615	10.0
22	10.719	0.622	11.6
8	2.281	0.629	8.4
27	8.778	0.656	9.2
20	4.580	0.668	8.9
10	3.062	0.685	8.8
32	5.060	0.686	8.6
7	2.949	0.722	8.6
19	8.320	0.732	9.1
31	15.250	0.746	8.9
36	12.118	0.766	8.0
24	5.667	0.771	8.0
12	2.000	0.776	8.3
11	2.520	0.799	8.1
23	6.120	0.804	8.4
35	21.000	0.812	8.2

was taken as the grand mean of 100 sample pattern matching scores obtained from each experimental structure-sample size combination. Using .80 as the target adjusted R^2 value, we first performed a linear regression on the smallest sample size combination (N=17) and incremented the sample sizes until reaching our target. If the largest sample size combination (N=500) failed to reach the target, the procedure was repeated with a Cobb-Douglas type function. In the case of factor pattern matching scores, a satisfactory linear regression model for each observed factor pattern was obtained with the smallest sample size. The independent variables in each model consisted of the stability ratio, uniqueness index, complexity index, all two variable interaction terms and the three variable interaction term. Table 4.2 summarizes the adjusted R^2, mean square error, entering variable coefficients and variable significance for each model.

According to the adjusted R^2 and mean square errors, the three models appear to predict matching scores reasonably well. The residual plots (Appendix E) also indicate a high degree of model aptness. Note in Table 4.1 that $Uniqeness^2$ is the second most significant independent variable in each model. However, in the model for unrotated matches, $Uniqueness^2$ is much less significant than in the models for our raw and normal varimax criterion. Further, the significance of this particular independent variable is nearly equal for the two rotation criterion. In fact, the raw and normal varimax models are very similar in terms of main and two-factor interaction effects. Also note that $Stability$ is nearly as significant as $Uniqueness^2$ in the unrotated model but very small relative to $Uniqueness^2$ in either rotated model. The coefficients of $Uniqueness^2$ and $Stability$ in each model also indicate that matching scores for rotated factor patterns are influenced by row effects (uniqueness) to a greater extent than unrotated factor patterns. Conversely, matching scores for unrotated factor patterns are influenced by column effects (stability) much more than the rotated factor patterns. Also shown are very large negative coefficients for $complexity$ in all three models. Evidently complexity in terms of geometry has a large

Table 4.2. Linear Regression Models For Factor Pattern Matching Score

	Variables	Coefficients	t-value
Unrotated	*Const*	11.707	26.96
	$Uniq^2$	-7.230	-6.70
	Stab	0.408	5.33
	Comp	-65.006	-3.05
	$Stab^2$	-0.005	-2.84
	$ST * Un$	-0.266	-2.57
	R^2 =.88	MSE =0.614	
Raw	*Const*	16.644	65.63
	$Uniq^2$	-13.014	-22.17
	Comp	-251.750	-4.61
	Stab	-0.032	-2.34
	$Comp^2$	5639.798	2.32
	R^2 =.95	MSE =0.396	
Normal	*Const*	16.675	65.32
	$Uniq^2$	-13.190	-22.43
	$Comp^2$	9619.153	3.17
	Comp	-237.73	-3.09
	Stab	-0.032	-2.37
	$St * Un * Co$	153.611	2.29
	St*C0	-115.069	-2.21
	R^2 =.95	MSE =0.386	

influence on factor pattern recovery. Since geometry can be considered a combination of the row and column effects that have already been found to be significant in each model, this makes intuitive sense. It is interesting that stability and uniqueness are present as an interaction term in the unrotated model. The results of this regression analysis seems to support our initial conclusions based upon the exploratory scatter plots.

4.2.3 Factor Analysis. A factor analysis was performed using the unrotated, raw, and normal factor scores as our observed variables. Three variables limited our dimensionality to three possible latent factors. Factor analysis was accomplished to determine if our initial conclusions about the relationship between pattern recovery and the three complexity effects were correct. If these conclusions are in error, factor analysis can help identify the presence of other effects not captured by our predictor variables. After extracting the common factors, the factor pattern was rotated using the normal varimax criterion. This rotated factor pattern is illustrated in Table 4.3.

Table 4.3. Rotated Factor Pattern For Matching Score

	f_1	f_2	f_3
Unrotated	.422	**.906**	.000
Raw	**.906**	.422	-.016
Normal	**.906**	.422	.016

The structure indicated by the bold faced factor loadings shows the observed matching scores for the raw and normal varimax criterion to be strongly influenced by the same underlying process. Observed matching scores for the unrotated factor pattern are influenced just as strongly by a different, independent process. It would be reasonable to infer that uniqueness is the underlying process associated with the rotated matching scores, while stability ratio is a similar underlying process for Unrotated matching scores. The two- factor plot (Appendix F) graphically illustrates this relationship. The cluster comprising group I are characterized by high stability ratios, while that of group II are characterized by very low average uniqueness.

4.2.4 Conclusions. Based upon the preceding analysis we can conclude That there is no difference between the ability of the raw or normal varimax rotation criterion to restore a known factor pattern (of the type we sampled from). Also, the difference between rotated and unrotated pattern recovery is negligible at high levels of average uniqueness in the patterns. Finally, we saw that stability ratio heavily influenced the unrotated factor pattern, while average uniqueness influenced the rotated factor patterns. In all cases, the number of pattern matches were predictable with a linear regression model at a very small sample size.

4.3 Factor Loading Residuals

In addition to pattern matching, our investigation also included each rotation criterion's ability to recover the factor loading of each experimental factor pattern. To address this performance area, we proposed using the root mean square residuals (RMSR) of the observed factor loadings as a measure of merit. RMSR was calculated for the permuted factor pattern that yielded the best pattern matching score.

4.3.1 Exploratory Scatter Plots. Figures 4.7, 4.8 and 4.9 show initial scatter plots for RMSR against each complexity effect. Unfortunately, none of these scatter plots indicate any conclusive relationship between factor loading recovery and the row, column or geometry effects. We can see that, regardless of the effect, there is very little difference between residuals of the observed factor patterns for the unrotated, raw varimax or normal varimax. The scatter plots for stability and uniqueness show a very slight trend in which the residuals decrease with increasing complexity in terms of column and row effects respectively. Intuitively, one would expect to see the opposite trend. A regression analysis was conducted to see if we could gain further insight into the underlying process affecting observed factor pattern loading residuals.

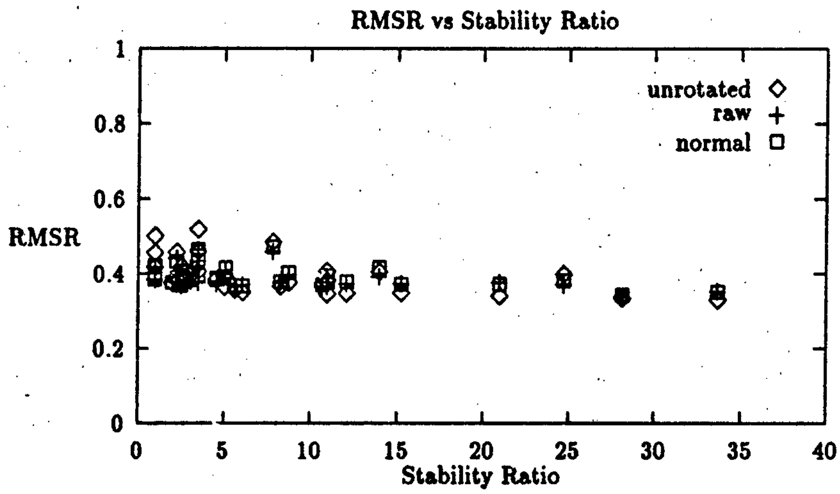


Figure 4.7. Stability Ratio vs Root Mean Square Loading Residuals For Small Sample Size N=17

4.3.2 Regression Analysis. As in the regression analysis conducted for pattern matching scores, RMSR was taken as the grand mean of 100 RMSR samples obtained from each experimental structure-sample size combination. Again using .80 as the target adjusted R^2 value, we first performed a linear regression on the smallest sample size combination and incremented the sample sizes until reaching the target value. If the largest sample size combination failed to reach the target, the procedure was repeated with a Cobb-Douglas type function. In the case of unrotated factor pattern RMSR, a satisfactory linear regression model was obtained with the smallest sample size. As for RMSR associated with the rotated criterion, we were unable to achieve the desired adjusted R^2 with the smallest sample size combination, and saw no improvement at any of the larger sample sizes. When attempting to fit a Cobb-Douglas function, we observed adjusted R^2 values lower than had been achieved with

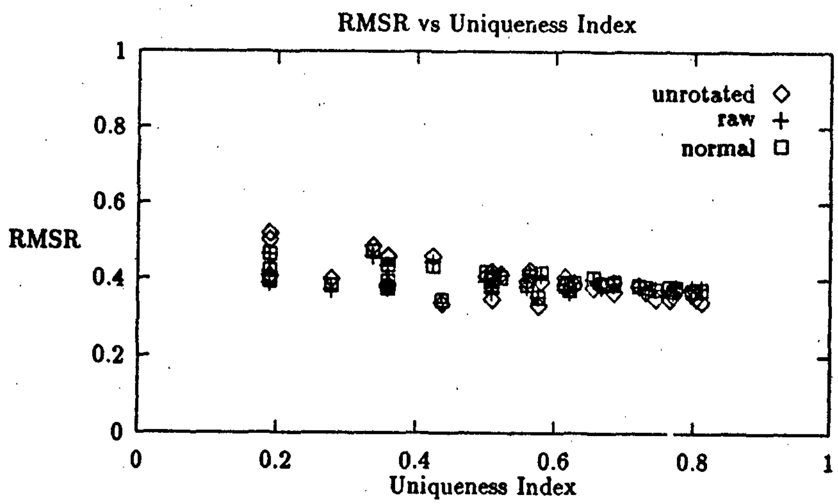


Figure 4.8. Uniqueness Index vs Root Mean Square Loading Residuals For Small Sample Size N=17

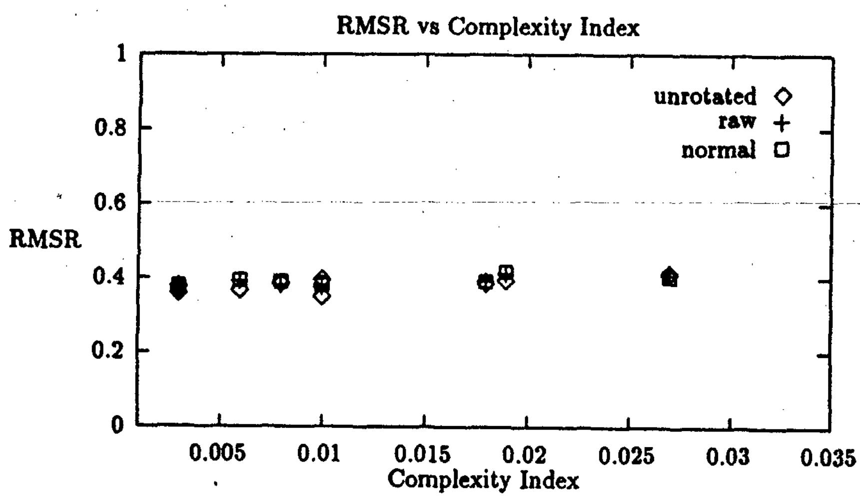


Figure 4.9. Complexity Index vs Root Mean Square Loading Residuals For Small Sample Size N=17

the linear model. Table 4.4 summarizes the adjusted R^2, mean square error, entering variable coefficients and variable significance for each linear model. The independent variables in each model consisted of the stability ratio, uniqueness index, complexity index, all two variable interaction terms and the three variable interaction term.

Table 4.4. Linear Regression Models For Root Mean Square Residuals

	Variables	Coefficients	t-value
Unrotated	$Const$	0.552	35.64
	$Uniq$	-0.238	-8.77
	$Stab$	-0.006	-4.71
	$Comp$	-65.006	-3.05
	$ST * Un$	0.006	2.66
	R^2 =.81	MSE =0.0003	
Raw	$Const$	0.468	30.57
	$Uniq$	-0.118	-4.40
	$Stab$	-0.004	-3.31
	$ST * Un$	0.005	-2.20
	R^2 =.54	MSE =0.0003	
Normal	$Const$	0.448	42.19
	$Uniq$	-0.087	-4.93
	$Stab^2$	-0.0001	-4.02
	R^2 =.53	MSE =0.0003	

According to the adjusted R^2 and mean square errors, the unrotated model appears to predict RMSR very well. The rotated models show only a fair level of RMSR prediction. Even so, the residual plots (Appendix G) indicate a high degree of aptness for all three models. Note in Table 4.4 that a *Uniqeness* term is again the second most significant independent variable in each model. However, in the model for unrotated matches, *Uniqueness* is twice as significant as in the models for our raw and normal varimax criterion. Perhaps this variable has a greater influence on the unrotated residuals, hence providing a better fit with the linear model. The significance of the *Stability* variable is approximately equal in each model. The coefficients of *Uniqueness* terms in each model indicate residuals for rotated factor patterns are influenced by row effects (uniqueness) to a greater extent

Table 4.5. Rotated Factor Pattern For Root Mean Square Residuals

	f_1	f_2	f_3
Unrotated	.497	**.866**	.033
Raw	**.865**	.496	-.073
Normal	**.836**	.518	.179

than unrotated factor patterns. This difference is only relative, since the influence in any specific model is very small. According to the coefficients for the *stability* terms, influence due to column effects (stability) is also very small. In every model, the coefficients of both variables have negative values. This supports our initial scatter plot observations where we detected a slight reduction in all residuals as uniqueness and stability increased.

4.3.3 Factor Analysis. A factor analysis was performed using the unrotated, raw and normal loading residuals as our observed variables. Three variables limited our dimensionality to three possible latent factors. Since the results of the scatter plots and regression analysis were largely inconclusive, a factor analysis was done to perhaps clarify the relationship underlying factor loading generation. After extracting the common factors, the factor pattern was rotated using the normal varimax criterion. This rotated factor pattern is illustrated in Table 4.5.

The structure indicated by the bold faced factor loadings shows the observed loading residuals for the raw and normal varimax criterion to be strongly influenced by the same underlying process. Observed residuals for the unrotated factor pattern are influenced just as strongly by a different, independent process. Recall that this is the same structure pattern identified in the factor analysis for factor pattern matching score. Notice, however, the significance of these loadings has decreased. A reduction in the influence of column and row effects has thus far been observed when comparing the exploratory plots and regression models of the loading residuals to those of the matching scores. As in the factor analysis for factor pattern matching score, it would be reasonable to infer that uniqueness is still the underlying process associated with the loading residuals of our rotated criterion, and stability ratio is again the underlying process for the unrotated loading residuals. It is interesting that while loadings on the first two factors decreased, we saw a third factor beginning to emerge. This may be in response to the geometry effect as measured by the complexity index. In retrospect, a broader range of experimental factor pattern complexity index might have answered this question. The corresponding factor plot (Appendix H). The absence of any noticeable clustering of factor scores would seem to indicate a complex relationship underlying the ability of unrotated and rotated varimax factor patterns to restore known experimental factor loadings

4.3.4 Conclusions. Although the preceding analysis was largely inconclusive ,we can still make some broad generalizations. First, according to our RMSR criterion there is little difference between the ability of the unrotated or either rotated factor pattern to restore the factor loadings of the experimental factor patterns. Second, all observed factor pattern loadings were influenced to a small extent by experimental factor pattern stability and uniqueness. Further, the uniqueness influence was opposite of what we expected. While there appeared to be an additional effect present, we were unable to conclude whether this was the effect due to geometry. Finally, a linear model can predict the loading residuals for the unrotated factor pattern fairly well, but only moderately so for the rotated factor patterns.

4.4 Eigenvalue Residuals

The final performance area explored in this study was the ability of the observed factor patterns to restore the eigenvalues of each experimental factor pattern. Recall from chapter 2, Kaiser's normal varimax rotation criterion had a leveling effect on the eigenvalues corresponding to the rotated factor pattern in that it reduced their divergence. This area was addressed similar to the factor pattern loading issue by using the root mean square eigenvalue residuals (RMSER) of the observed factor loadings as a measure of merit. In the interest of consistency, RMSER was calculated for the same permuted factor pattern that yielded the best pattern matching score.

4.4.1 Exploratory Scatter Plots. Figures 4.10, 4.11 and 4.12 show initial scatter plots for RMSER against the column row and geometry complexity effects. Obviously, the relationship between eigenvalue recovery and our chosen complexity effects appears to be very complicated. As illustrated in the scatter plots, there is no readily apparent, consistent residual response with respect to experimental factor pattern stability, uniqueness or complexity. It is also tempting to conclude in general that there is no substantial difference between the ability of the unrotated, or either rotated factor patterns to recover the eigenvalues of the experimental factor patterns. However, the apparent complexity of the relationship between our measured effects and eigenvalue residuals forces us to withhold this conclusion. A regression analysis may simplify this relationship.

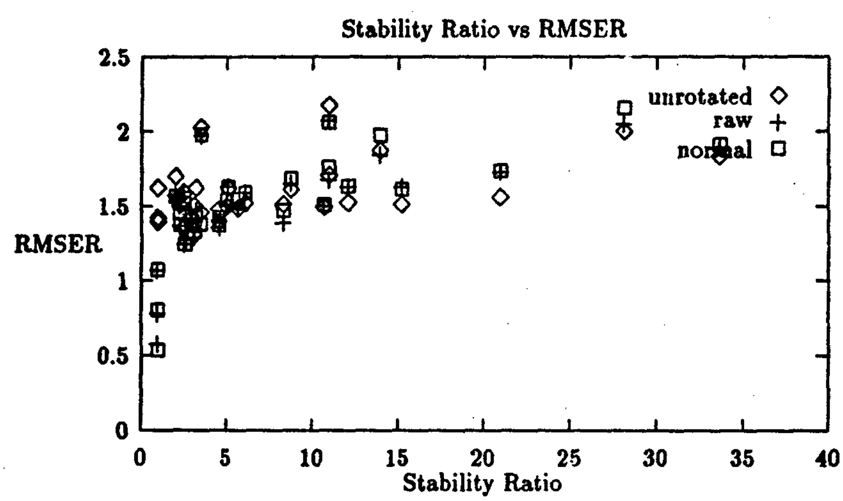


Figure 4.10. Stability Ratio vs Root Mean Square Eigenvalue Residuals Sample Size N=17

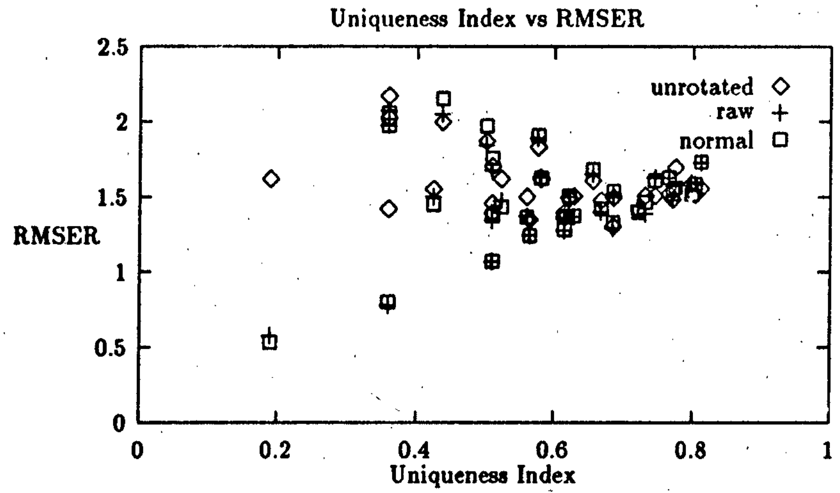


Figure 4.11. Uniqueness Index vs Root Mean Square Eigenvalue Residuals Sample Size N=17

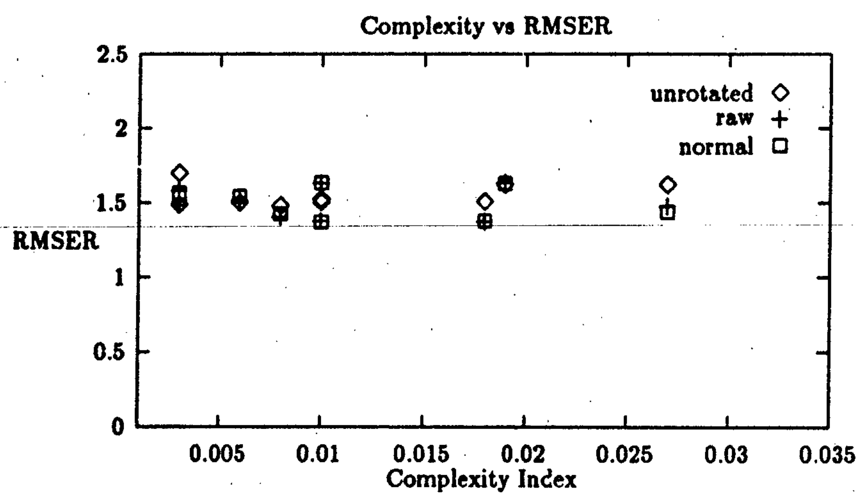


Figure 4.12. Complexity Index vs Root Mean Square Eigenvalue Residuals Sample Size N=17

4.4.2 Regression Analysis. Consistent with all previous regression analysis, RMSER was taken as the grand mean of 100 sample RMSER samples obtained from each experimental structure-sample size combination. Still using .80 as the target adjusted R^2 value, we first performed a linear regression on the smallest sample size combination and incremented the sample sizes until reaching the target value. If the largest sample size combination failed to reach the target, the procedure was repeated with a Cobb-Douglas type function. For the unrotated and both rotated factor patterns, we were able to satisfactorily model RMSER with a linear regression model using the smallest structure-sample size combinations. Table 4.6 summarizes the adjusted R^2, mean square error, entering variable coefficients and variable significance for each linear model. The independent variables in each model again consisted of the stability ratio, uniqueness index, complexity index, all two variable interaction terms and the three variable interaction term.

Table 4.6. Linear Regression Models For Root Mean Square Eigenvalue Residuals

	Variables	Coefficients	t-value
Unrotated	$Stab$	0.344	8.10
	$St * Un$	-0.352	-6.26
	$Constant$	1.515	4.28
	$Stab^2$	-0.002	-2.62
	$Unique$	-1.351	-2.23
	R^2 =.88	MSE =0.1859	
Raw	$St * Un$	-0.415	-10.59
	$Stab$	-0.364	9.26
	$Constant$	0.737	4.26
	$Stab^2$	-0.002	-2.01
	R^2 =.84	MSE =0.2054	
Raw	$Stab$	-0.300	13.36
	$St * Un$	-0.407	-10.33
	$Constant$	0.916	8.34
	R^2 =.84	MSE =0.2092	

We can readily see from the values for adjusted R^2 and mean square error that all models explain most of the variance of the eigenvalue residuals. The corresponding residual plots (Appendix I) indicate we have an apt model. After accounting for the relative small scale of the residual axis and disregarding the obvious outliers, we see that the variance of the residuals is fairly constant. A closer look at the entering variables for each model reveals several interesting observations. Each model incorporates similar independent variables and these variables are mostly interaction terms. This explains why our exploratory factor plots were so difficult to interpret. Note that *Stability* or the interaction of *Stability* with *Uniqueness* are the most significant variables in each model. In addition, it is interesting that the coefficients of each term are nearly equal in value but opposite in sign. Also note that no form of the complexity index appears in any model. Despite these successful models, the influence of column, row and geometry effects is still not clear.

4.4.3 Factor Analysis. A factor analysis was performed using the unrotated, raw and normal eigenvalue residuals as our observed variables. Three variables limited our dimensionality to three factors. This analysis should determine if a simple structure is present in the residual eigenvalue data. After extracting the common factors, the factor pattern was rotated using the normal varimax criterion. This rotated factor pattern is illustrated in Table 4.7.

Table 4.7. Rotated Factor Pattern For Root Mean Square Residuals

	f_1	f_2	f_3
Unrotated	.623	**.761**	.177
Raw	**.689**	.674	.264
Normal	**.761**	.624	.172

The structure indicated by the bold faced factor loadings is identical to that of the factor analysis for matching score and loading residuals. However, in this case we have a very complicated structure with near equal moderate loadings on adjacent factors. Note that our third factor loadings are becoming more significant relative to the first two factors. Based upon this structure and the many interaction terms in the preceding regression analysis, we can not identify any meaningful, underlying relationship for the eigenvalue residual data. The absence of a simple structure indicates our current metrics for measuring performance in terms of eigenvalue residuals are insufficient. The factor plot for this analysis provides no additional insight into the data structure, but it is included (Appendix J) for completeness. The absence of any noticeable clustering of factor scores further illustrates the complexity of this particular data structure

4.4.4 Conclusions. Although the eigenvalue residuals were predictable with a linear regression model, these functions were very complicated and difficult to interpret. Perhaps a different set of independent variables would enable us to identify a simpler underlying structure. Complexity due to geometry may still have some influence on eigenvalue residuals, but our narrow complexity index range does not allow us to assess this possibility.

4.5 Summary

At the onset of this study we intended to investigate three issues with respect to Kaiser's normal varimax criterion:

- How well does the normal varimax perform over a wide range of sample size, structure complexity, variable communality and factor loading?
- In addition to restoring the known structure of an experimental factor pattern, how effective is normal varimax in terms of returning variable loadings and eigenvalues?
- With what types of data will normal varimax not outperform raw varimax?
- Are there circumstances where neither rotation criterion is desirable?

Some of our measures of merit assessed these performance areas very well and others were not so successful. In particular, pattern matching in conjunction with our chosen predictor variables best described the response of pattern recovery and how it is influenced by different aspects of factor pattern complexity. Although RMSR and RMSER were very predictable, they did not allow us to identify a simple underlying relationship with respect to factor pattern loading and eigenvalue recovery. We propose that perhaps these are not key performance areas and factor pattern structure recovery should be considered of primary importance. In terms of factor pattern recovery, we clearly saw that the raw and normal varimax criterion were both influenced by factor pattern uniqueness and stability. At no time did we see normal varimax outperform raw varimax in any performance area. Furthermore we saw that at high levels of factor pattern uniqueness, neither rotated factor pattern outperformed the unrotated factor pattern in factor pattern recovery. With more study this final observation may provide valuable insight to the researcher conducting factor analysis. If during his analysis he finds no difference between the rotated and unrotated factor patterns, perhaps the researcher can conclude that he already has already identified the simple structure, or there is no simple underlying structure present in his data. At that point he should select a different set of measured variables, extract a factor pattern, and rotate to determine which is the case.

V. Recommendations

5.1 Future Research

Future studies on the sensitivity of factor analysis techniques should certainly be pursued. The source code produced for this Monte Carlo study is well suited to work beyond the scope of this effort. The idea of geometric complexity may be valid, but the range of our complexity index was too narrow to really asses the validity of this assumption. Future work should also address other methods for characterizing factor pattern behavior. Additional performance areas must also be addressed before concluding factor pattern recovery is sufficient for assessing rotation criterion accuracy. Obviously, future efforts should also be expanded to include rotation methods other than varimax criterion.

We focused on smaller sample sizes to explore the threshold for satisfactory factor rotation. As measured by factor pattern recovery, this threshold appears to be lower than expected. Perhaps RMSR and RMSER are heavily influenced by sample size. This should be answered. In addition, sample size should be investigated for its effect on rotation reliability where reliability is determined by the standard deviation of factor pattern scores.

In this effort we made a dimensionality assumption by retaining four factors. This was done to be consistent with Kaiser's dimensionality in his validation of the normal varimax criterion. Different variable-factor ratios may have a significant impact on factor rotation.

Recall that when calculating the pattern matching scores, we converted each observed factor pattern to ones and zeros. This provided a simple method for calculating matching scores, but some information was lost. Two adjacent factor loading may be very close in value, yet one will ultimately be changed to a zero loading. The zero loading may have been the true significant loading in the experimental factor pattern and just missed being matched. These types of occurrences also need to be considered.

Finally, the Monte Carlo method is a very valuable technique that should be applied to research outside the field of factor analysis. Modern personal computers and work stations make this approach readily accessible.

Appendix A. Data Set Generation

A.1 FORTRAN code

This Appendix lists the FORTRAN code for generating the sample data. It takes a covariance matrix, performs a Choleski decomposition and generates 100 data sets of random multivariate normal deviates for experimental factor pattern number one (in this case), smallest (N=17) sample size. The data file is formatted as a 1700*4 matrix.

```
      PROGRAM RANDOM
INTEGER    I, IRANK,  ISEED,  J, K, LDR, LDRSIG, NOUT, NR, N
REAL      R(17,16), RSIG(16,16),RCOV(16,16)

EXTERNAL  CHFAC, RNMVN, RNSET, UMACH

        CALL UMACH(2,NOUT)
        NR=17
        K =16
        LDRSIG=16
        LDR=17

*****OPEN THE FILE CONTAINING THE COVARIANCE MATRIX FOR EXPERIMENTAL****
*****FACTOR PATTERN 1, CALCULATED FROM THE PRODUCT OF THE LOADING*******
*****MATRIX AND ITS TRANSPOSE.  DONT FORGET TO HAVE ADDED THE DIAGONAL**
*****MATRIX OF THE UNIQUENESS VALUES.  READ COVARIANCE MATRIX AND OPEN**
*****FILE THAT WILL HOLD 100 SAMPLE SETS OF RANDOM MULTIVARIATE NORMAL**
*****DEVIATES***********************************************************

*****THE LEADING 6*6 SUB MATRIX OF THE COVARIANCE MATRIX IS SHOWN BELOW*
*
*        1      .81      .81      .81       0        0
*      .81        1      .81      .81       0        0
*      .81      .81        1      .81       0        0
*      .81      .81      .81        1       0        0
*        0        0        0        0       1      .81
*        0        0        0        0     .81        1

        OPEN(3,FILE='structure.dat')
        READ(3,*),  ((RCOV(I,J),J=1,K),I=1,K)
        CLOSE(3)
        OPEN(1,FILE='sample.dat')

*****CREATE DATA SETS INCREMENTING THE SEED WITH EACH SET****************

        ISEED = 123457
        DO 20 N=1,100
        CALL CHFAC  (K, RCOV, 16, 0.00001, IRANK, RSIG, LDRSIG)
        CALL RNSET(ISEED)
        CALL RNMVN (NR, K, RSIG, LDRSIG, R, LDR)

*****WRITE THE 100 DATA SETS TO OUTPUT FILE*****************************

        WRITE (1,10) ((R(I,J),J=1,K),I=1,NR)
        ISEED = ISEED+100
20      CONTINUE
10      FORMAT (1X,16F8.4)

        CLOSE(1)
        END
```

```
*****THE LEADING 6*6 SUB MATRIX OF THE 16*1700 DATA SET IS SHOWN BELOW***
*
*   2.0516  1.5520  1.5405  2.0164  0.4144  0.1419
*   1.0833  1.1174 -0.0433  1.7127  1.2355  1.9269
*   0.0826  0.6900  0.1764  1.3174  0.9498  1.0576
*   1.2777  0.5152  0.1561  0.2775  0.9056  1.0729
*  -1.2260 -0.7603 -0.6753 -0.6706  0.8173  0.6975
*   0.3385  0.4181  1.8470  1.1316  0.8911  1.3856
*
```

Appendix B. Index Routine

B.1 FORTRAN code

This Appendix lists the FORTRAN code for creating an index vector for sub-grouping the sample data file into 100 data sets having the appropriate number of observations.

```
      PROGRAM INDEX
INTEGER   K, NR, N, I, L, E(16,1)
        NR      =16
        K       =17
        T       =K*NR
        OPEN(18,FILE='index.dat',FORM='FORMATTED')
        I=0
        DO 200 L=1,100
        DO 100 N=1,17
           E(K,1)=I+1
           WRITE (18,10) E(K,1)
100     CONTINUE
        I=I+1
200     CONTINUE
10      FORMAT (I3)
        CLOSE(18)
        END
```

Appendix C. Unrotated Factor Pattern Generation

C.1 SAS code

This Appendix lists the SAS code for gererating the observed unrotated factor pattern for each of the 100 sampled data sets. The data file produced by the data generation code is merged with an index vector so we can factor analyze 100 subgroups corresponding to the appropriate sample size. For example, the data file for the small sample size will be a 1700*16 matrix. The index vector will range from 1 to 100 with 17 iterations of each number. SAS will factor analyze the first 17 observations, second 17 observations and so forth, until it has factor analyzed 100 sets of 17 observations.

```
options linesize=78;
FILENAME NEW 'sample.dat';
DATA XX1;
 INFILE NEW;
  input X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16;

DATA XX2;
 INFILE 'index.dat';
 input index;
DATA XX3;
merge XX2 XX1;
proc factor data=XX3 method=prin nfactors=4;
by index;
```

```
---------------------------------- INDEX=1 ----------------------------------
Initial Factor Method: Principal Components

                        Prior Communality Estimates: ONE

      Eigenvalues of the Correlation Matrix:  Total = 16  Average = 1

                            1            2            3            4
      Eigenvalue       6.2965       3.1123       2.8689       1.9385
      Difference       3.1842       0.2434       0.9304       1.4314
      Proportion       0.3935       0.1945       0.1793       0.1212
      Cumulative       0.3935       0.5881       0.7674       0.8885

                            5            6            7            8
      Eigenvalue       0.5071       0.4127       0.2698       0.2339
      Difference       0.0944       0.1429       0.0359       0.1125
      Proportion       0.0317       0.0258       0.0169       0.0146
      Cumulative       0.9202       0.9460       0.9629       0.9775

                            9           10           11           12
      Eigenvalue       0.1213       0.1008       0.0560       0.0465
      Difference       0.0205       0.0448       0.0095       0.0279
      Proportion       0.0076       0.0063       0.0035       0.0029
      Cumulative       0.9851       0.9914       0.9949       0.9978

                           13           14           15           16
```

Eigenvalue	0.0187	0.0143	0.0027	0.0001
Difference	0.0044	0.0115	0.0027	
Proportion	0.0012	0.0009	0.0002	0.0000
Cumulative	0.9989	0.9998	1.0000	1.0000

4 factors will be retained by the NFACTOR criteri...

Factor Pattern

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
X1	-0.53209	0.73981	0.14878	0.09028
X2	-0.50755	0.74434	0.02454	0.30341
X3	-0.55022	0.73031	0.16927	-0.07489
X4	-0.32998	0.82213	0.28371	0.11522
X5	0.63843	0.20186	0.39150	-0.49487
X6	0.64913	0.48773	0.23666	-0.47719
X7	0.48365	-0.12627	0.65814	-0.46377
X8	0.64715	0.26590	0.36936	-0.39919
X9	-0.48783	-0.17746	0.73005	0.12092
X10	-0.59957	-0.34110	0.65957	0.11973
X11	-0.64469	-0.30897	0.58889	-0.02810
X12	-0.70671	-0.33157	0.56528	0.07712
X13	0.70811	0.27420	0.14132	0.51245
X14	0.76861	-0.00551	0.31612	0.51030
X15	0.82272	-0.07591	0.30864	0.43758
X16	0.76217	0.04723	0.32964	0.47136

Appendix D. Monte Carlo Routine

D.1 FORTRAN code

This Appendix lists the FORTRAN code for the Monte Carlo method used in this research. The program reads the experimental factor patterns and the SAS generated unrotated factor patterns. It utilizes a loop to read 100 observed unrotated factor patterns that have been stored in one file by reading the first 1-16*4 factor loadings, the second 17-32*4 factor loadings ect. It then calculates all predictor values, performs the raw and normal varimax rotations and calculates all measures of merit. The measures of merit are taken at the grand mean of the 100 observed foactor patterns of each type (unrotated, raw and normal) produced from every experimental factor pattern.

```
      INTEGER    LDA, LDB, LDT, MAXIT, NF, NORM, NVAR, K,
     &           F, ROW, COL, SAMPLE, S, STR, Z,
     &           SUMI, SUMP, SUMIR, SUMPR, SUMIK, SUMPK,
     &           HUN, HRW, HKI, COUNT

      REAL       EPS, W, SR, ESR,PESR,AMESR,
     &           SESR, RSESR, KSESR, AIRHITS, ABRHITS,
     &           SSR, RSSR, KSSR, AIIHITS, AIBHITS,
     &           AMSR, AIKHITS, ABKHITS, RAMESR, RAMSR,
     &           KAMSR, TOTSIROW, STRAT, HINDEX, PAMESR,
     &           CINDEX, THINDEX, PAMESR, PKAMESR, PRAMESR,
     &           PSESR, PRSESR,PAMSR,PESR,
     &           PESR, USSE, SDUM, USSM, SDUM, RSSE, SDRE,
     &           RSSM, SDRM, KSSE, SDKE, KSSM, SDKM ,UUSSE,
     &           SDUE, SDUUE,PUSSM,PSDUM,URSSE,USDRE,PRAMSR,
     &           PRSSM,PSDRM,KAMESR,UKSSE,USDKE,PKSSM,PSDKM


      PARAMETER  (EPS=0.0, LDA=16, LDB=16, LDT=4, MAXIT=30,
     &            NF=4, NORM=K, NVAR=16, W=1.0, LDD=160,
     &            SAMPLE=100)

      REAL       A(LDA,NF), B(LDB,NF), T(LDT,NF),
     &           LOAD(NVAR,NF), C(16,4),
     &           PRNUM(16,4,SAMPLE), PNNUM(1,4,SAMPLE),
     &           RNUM(16,4,SAMPLE), UNIQ(16), PKNUM(1,4,SAMPLE),
     &           TOTH(16),NNUM(1,4,SAMPLE), KNUM(1,4,SAMPLE), INUM(4),
     &           R(16,4,SAMPLE), R(16,4),ER(1,4,SAMPLE),
     &           PERM(16,4,24), PER(1,4,SAMPLE ), PRER(1,4,SAMPLE),
     &           PMESR(SAMPLE), MSR(SAMPLE), PKMESR(SAMPLE),
     &           KMSR(SAMPLE), PRMESR(SAMPLE), RMSR(SAMPLE),
     &           MESR(SAMPLE), PER(1,4,SAMPLE),PMSR(SAMPLE),
     &           RER(1,4,SAMPLE), RMESR(SAMPLE), PRMESR(SAMPLE),
     &           PRMSR(SAMPLE),KMESR(SAMPLE),PKMSR(SAMPLE),
     &           CB(16,4),CK(16,4)

      INTEGER     D(16,4), TRLOAD(4,16), PRLOAD(4,4,24),
```

```
     &          NUMHITS(24,SAMPLE),U(16,4,24), RV(16,4,24),
     &          KV(16,4,24)

      EXTERNAL   FROTA, WRRRN

      F=0

******CREATE THE DATA FILE ALL STATISTICAL OUTPUT******************

      OPEN( UNIT=26, FILE='out.dat')

******OPEN AND READ THE DATA FILE CONTAINING THE EXPERIMENTAL******
******LOADINGS MATRIX**********************************************

      OPEN(1,FILE='in.dat')
      READ(1,*), ((LOAD(I,J),J=1,NF),I=1,NVAR)
      CLOSE(1)

******CALCULATE THE STURCTURE INDEX FOR THE EXPERIMENTAL LOADINGS**
******MATRIX*******************************************************

      CINDEX=0.0
      TOTSIROW=0.0
      DO 822 I=1,16
      TOTSIROW=TOTSIROW+((LOAD(I,1)*LOAD(I,2))**2+(LOAD(I,1)*LOAD(I,3))**2
     & +(LOAD(I,1)*LOAD(I,4))**2+(LOAD(I,2)*LOAD(I,3))**2
     & +(LOAD(I,2)*LOAD(I,4))**2+(LOAD(I,3)*LOAD(I,4))**2)
822   CONTINUE
      CINDEX=TOTSIROW/16.0

******CALCULATE THE UNIQUENESS INDEX FOR THE EXPERIMENTAL LOADINGS**
******MATRIX**********************************************************

      THINDEX=0.0
      DO 932 I=1,16
         UNIQ(I)=0.0
         TOTH(I)=0.0
932   CONTINUE

      DO 842 I=1,16
      DO 832 J=1,4
         TOTH(I)=TOTH(I)+(LOAD(I,J)**2)
832   CONTINUE

      UNIQ(I)=1-TOTH(I)
      THINDEX=THINDEX+UNIQ(I)
842   CONTINUE
      HINDEX=THINDEX/16.0

******CONDUCT BINARY TRANSFORMATION OF EXPERIMENTAL LOADINGS MATRIX***

      DO 1021 I=1,NVAR
      DO 1020 J=1,NF
         D(I,J)=0
1020  CONTINUE
1021  CONTINUE

      DO 1032 I=1,NVAR
         BIG=ABS(LOAD(I,1))
         LOCBIG=1
         D(I,1)=1
      DO 1031 J=2,NF
         IF (ABS(LOAD(I,J)) .GT. BIG) THEN
            D(I,J)=1
            D(I,LOCBIG)=0
            BIG=ABS(LOAD(I,J))
            LOCBIG=J
         END IF
1031  CONTINUE
1032  CONTINUE

******IF DESIRED, LOADINGS MATRIX OF 1'S AND O'S CAN BE WRITTEN OUT***
*       PRINT*,'SIMPLE STRUCTURE PATTERN'
```

```
*       PRINT*
*       WRITE(*,14)((D(I,J),J=1,NF),I=1,16)
*       PRINT*

*       WRITE(26,*)' '

*       WRITE(26,7000)
*7000   FORMAT(21X,'INITIAL FACTOR PATTERN')

*       WRITE(26,*)' '

*       WRITE(26,7001) ((D(I,J),J=1,NF),I=1,16)
*7001   FORMAT(22X,4I4)

******CALCULATE THE TRANSPOSE OF THE INITIAL STRUCTURE SO PATTERN*****
******MATCHING SCORE CAN BE CALCULATED*********************************

      COL=1
      DO 1132 I=1,NF
         ROW=1
      DO 1131 J=1,NVAR
         TRLOAD(I,J)=D(ROW,COL)
         ROW=ROW+1
1131  CONTINUE

      COL=COL+1
1132  CONTINUE

******CALCULATE EIGENVALUES OF INITIAL LOADING STRUCTURE**************

       DO 900 J=1,NF
          INUM(J)=0.0
       DO 800 I=1,NVAR
          INUM(J)=INUM(J)+LOAD(I,J)**2
800    CONTINUE
900    CONTINUE

*******CALCULATE STABILITY RATIO**************************************

       STRAT=0.0
       STRAT=STRAT+ABS(INUM(1)/INUM(4))

******AT THIS POINT, ALL THE EXPERIMENTAL FACTOR PATTERN HAS BEEN*****
******AND ALL PREDICTOR VARIABLES HAVE BEEN CALCULATED****************


******INITIATE LOOP***************************************************
******HERE WE INITIATE THE LOOP FOR READING IN 100 SAMPLE UNROTATED****
******FACTOR PATTERNS GENERATED IN SAS, CALCULATE OUR MEASURES********
******MERIT, AND STATISTICS*******************************************
******WE MUST ALSO INITIALIZE ALL OF OUR STATISTICS*******************

      SUMP=0
      SUMI=0
      SUMPR=0
      SUMIR=0
      SUMPK=0
      SUMIK=0
      HUN=0
      HRW=0
      HKI=0
      SESR=0.0
      SSR=0.0
      RSESR=0.0
      RSSR=0.0
      KSESR=0.0
      KSSR=0.0
      PSESR=0.0
      PRSESR=0.0
      PKSESR=0.0

******OPEN FILE AND READ THE 100 UNROTATED SAMPLE FACTOR PATTERNS*****
```

```
******GENERATED FROM EACH OF OUR Z=36 EXPERIMENTAL FACTOR PATTERNS****

      OPEN(2,FILE=INITIAL(Z))

      DO 300 S=1,100

      READ(2,*), ((A(I,J),J=1,NF),I=1,NVAR)

******(A(I,J) IS OUR SAMPLE UNROTATED FACTOR PATTERN******************

******GENERATE PERMUTATIONS OF OUR SAMPLE UNROTATED FACTOR PATTERN****
******IN PREPARATION FOR DETERMINING BEST PATTER MATCHING SCORE*******

      DO 2001 I=1,NVAR
      DO 2000 J=1,NF
      PERM(I,J,1)=A(I,J)
2000  CONTINUE
2001  CONTINUE

      DO 2002 I=1,16
      PERM(I,1,2)=A(I,1)
      PERM(I,2,2)=A(I,2)
      PERM(I,3,2)=A(I,4)
      PERM(I,4,2)=A(I,3)
2002  CONTINUE

      DO 2003 I=1,16
      PERM(I,1,3)=A(I,1)
      PERM(I,2,3)=A(I,3)
      PERM(I,3,3)=A(I,2)
      PERM(I,4,3)=A(I,4)
2003  CONTINUE

      DO 2004 I=1,16
      PERM(I,1,4)=A(I,1)
      PERM(I,2,4)=A(I,3)
      PERM(I,3,4)=A(I,4)
      PERM(I,4,4)=A(I,2)
2004  CONTINUE

      DO 2005 I=1,16
      PERM(I,1,5)=A(I,1)
      PERM(I,2,5)=A(I,4)
      PERM(I,3,5)=A(I,2)
      PERM(I,4,5)=A(I,3)
2005  CONTINUE

      DO 2006 I=1,16
      PERM(I,1,6)=A(I,1)
      PERM(I,2,6)=A(I,4)
      PERM(I,3,6)=A(I,3)
      PERM(I,4,6)=A(I,2)
2006  CONTINUE

      DO 2007 I=1,16
      PERM(I,1,7)=A(I,2)
      PERM(I,2,7)=A(I,1)
      PERM(I,3,7)=A(I,3)
      PERM(I,4,7)=A(I,4)
2007  CONTINUE

      DO 2008 I=1,16
      PERM(I,1,8)=A(I,2)
      PERM(I,2,8)=A(I,1)
      PERM(I,3,8)=A(I,4)
      PERM(I,4,8)=A(I,3)
2008  CONTINUE

      DO 2009 I=1,16
      PERM(I,1,9)=A(I,2)
      PERM(I,2,9)=A(I,3)
      PERM(I,3,9)=A(I,1)
      PERM(I,4,9)=A(I,4)
```

```
2009  CONTINUE
      DO 2010 I=1,16
      PERM(I,1,10)=A(I,2)
      PERM(I,2,10)=A(I,3)
      PERM(I,3,10)=A(I,4)
      PERM(I,4,10)=A(I,1)
2010  CONTINUE
      DO 2011 I=1,16
      PERM(I,1,11)=A(I,2)
      PERM(I,2,11)=A(I,4)
      PERM(I,3,11)=A(I,1)
      PERM(I,4,11)=A(I,3)
2011  CONTINUE
      DO 2012 I=1,16
      PERM(I,1,12)=A(I,2)
      PERM(I,2,12)=A(I,4)
      PERM(I,3,12)=A(I,3)
      PERM(I,4,12)=A(I,1)
2012  CONTINUE
      DO 2013 I=1,16
      PERM(I,1,13)=A(I,3)
      PERM(I,2,13)=A(I,1)
      PERM(I,3,13)=A(I,2)
      PERM(I,4,13)=A(I,4)
2013  CONTINUE
      DO 2014 I=1,16
      PERM(I,1,14)=A(I,3)
      PERM(I,2,14)=A(I,1)
      PERM(I,3,14)=A(I,4)
      PERM(I,4,14)=A(I,2)
2014  CONTINUE
      DO 2015 I=1,16
      PERM(I,1,15)=A(I,3)
      PERM(I,2,15)=A(I,2)
      PERM(I,3,15)=A(I,1)
      PERM(I,4,15)=A(I,4)
2015  CONTINUE
      DO 2016 I=1,16
      PERM(I,1,16)=A(I,3)
      PERM(I,2,16)=A(I,2)
      PERM(I,3,16)=A(I,4)
      PERM(I,4,16)=A(I,1)
2016  CONTINUE
      DO 2017 I=1,16
      PERM(I,1,17)=A(I,3)
      PERM(I,2,17)=A(I,4)
      PERM(I,3,17)=A(I,1)
      PERM(I,4,17)=A(I,2)
2017  CONTINUE
      DO 2018 I=1,16
      PERM(I,1,18)=A(I,3)
      PERM(I,2,18)=A(I,4)
      PERM(I,3,18)=A(I,2)
      PERM(I,4,18)=A(I,1)
2018  CONTINUE
      DO 2019 I=1,16
      PERM(I,1,19)=A(I,4)
      PERM(I,2,19)=A(I,1)
      PERM(I,3,19)=A(I,2)
```

```
      PERM(I,4,19)=A(I,3)
2019  CONTINUE

      DO 2020 I=1,16
      PERM(I,1,20)=A(I,4)
      PERM(I,2,20)=A(I,1)
      PERM(I,3,20)=A(I,3)
      PERM(I,4,20)=A(I,2)
2020  CONTINUE

      DO 2021 I=1,16
      PERM(I,1,21)=A(I,4)
      PERM(I,2,21)=A(I,2)
      PERM(I,3,21)=A(I,1)
      PERM(I,4,21)=A(I,3)
2021  CONTINUE

      DO 2022 I=1,16
      PERM(I,1,22)=A(I,4)
      PERM(I,2,22)=A(I,2)
      PERM(I,3,22)=A(I,3)
      PERM(I,4,22)=A(I,1)
2022  CONTINUE

      DO 2023 I=1,16
      PERM(I,1,23)=A(I,4)
      PERM(I,2,23)=A(I,3)
      PERM(I,3,23)=A(I,1)
      PERM(I,4,23)=A(I,2)
2023  CONTINUE

      DO 2024 I=1,16
      PERM(I,1,24)=A(I,4)
      PERM(I,2,24)=A(I,3)
      PERM(I,3,24)=A(I,2)
      PERM(I,4,24)=A(I,1)
2024  CONTINUE

******PUT ALL PERMUTTED MATRICES INTO BINARY FORM IN PREPARATION TO****
******POST MULTIPLY THE TRANSPOSE OF THE EXPERIMENTAL FACTOR PATTERN***
******TO DETERMINE BEST FACTOR PATTERN SCORE***************************

      DO 1093 H=1,24
      DO 1081 I=1,NVAR
      DO 1080 J=1,NF
         U(I,J,H)=0
1080  CONTINUE
1081  CONTINUE

      DO 1092 I=1,NVAR
         BIG=ABS(PERM(I,1,H))
         LOCBIG=1
         U(I,1,H)=1

      DO 1091 J=2,NF
      IF (ABS(PERM(I,J,H)) .GT. BIG) THEN
         U(I,J,H)=1
         U(I,LOCBIG,H)=0
         BIG=ABS(PERM(I,J,H))
         LOCBIG=J
      END IF
1091  CONTINUE
1092  CONTINUE
1093  CONTINUE

******CALCULATE PATTERN MATCHING SCORE FOR EACH PERMUTATION OF SAMPLE***
******OBSERVED UNROTATED FACTOR PATTERN*********************************
```

```
      DO 3000 H=1,24
      DO 1142 I=1,4
         PRLOAD(I,1,H)=0
      DO 1141 J=1,NVAR
         PRLOAD(I,1,H)=PRLOAD(I,1,H)+(TRLOAD(I,J)*U(J,1,H))
1141  CONTINUE
1142  CONTINUE

      DO 1144 I=1,4
         PRLOAD(I,2,H)=0
      DO 1143 J=1,NVAR
         PRLOAD(I,2,H)=PRLOAD(I,2,H)+(TRLOAD(I,J)*U(J,2,H))
1143  CONTINUE
1144  CONTINUE

      DO 1146 I=1,4
         PRLOAD(I,3,H)=0
      DO 1145 J=1,NVAR
         PRLOAD(I,3,H)=PRLOAD(I,3,H)+(TRLOAD(I,J)*U(J,3,H))
1145  CONTINUE
1146  CONTINUE

      DO 1148 I=1,4
         PRLOAD(I,4,H)=0
      DO 1147 J=1,NVAR
         PRLOAD(I,4,H)=PRLOAD(I,4,H)+(TRLOAD(I,J)*U(J,4,H))
1147  CONTINUE
1148  CONTINUE

      NUMHITS(H,S)=0
      DO 1149 I=1,4
      NUMHITS(H,S)=NUMHITS(H,S)+PRLOAD(I,I,H)
1149  CONTINUE

3000  CONTINUE

******DETERMINE BEST PERMUTATION BY DETERMINING HIGHEST SCORING*****
******(NUMHITS) OF ALL PERMUTATIONS***********************************

      NUN=NUMHITS(1,S)
      ID=0
      DO 3500 H=2,24
         IF (NUMHITS(H,S) .GT. NUN) THEN
            NUN=NUMHITS(H,S)
            ID=H
         END IF
3500  CONTINUE

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******UNROTATED FACTOR PATTERNS, CALCULATE THE AVERAGE**********
*******NUMBER OF MATCHES FOR THE SAMPLE***************************

      SUMP=SUMP+NUN
      IF (S .EQ. SAMPLE) THEN
         AIBHITS=FLOAT(SUMP)/FLOAT(SAMPLE)
      END IF

******CALCULATE EIGENVALUES BEST PERMUTTED FACTOR PATTERN********

      DO 275 J=1,4
         PRNUM(1,J,S)=0.0
      DO 250 I=1,16
         PRNUM(1,J,S) =PRNUM(1,J,S)+PERM(I,J,ID)**2
250   CONTINUE
275   CONTINUE

******CALCULATE THE EIGENVALUE RMSR FOR EACH BEST PERMUTED SAMPLE**
****** UNROTATED SAMPLE FACTOR PATTERN******************************

      DO 421 J=1,4
```

```
        PER(1,J,S)=(PRNUM(1,J,S)-INUM(J))**2
421   CONTINUE

      PESR=0.0
         DO 427 J=1,4
         PESR=PESR+PER(1,J,S)
427   CONTINUE

      PMESR(S)=SQRT(PESR/4.0)

******CALCULATE THE LOADING RMSR FOR EACH BEST PERMUTED SAMPLE*****
****** UNROTATED SAMPLE FACTOR PATTERN******************************

      DO 134 I=1,16
      DO 133 J=1,4
         R(I,J)=(LOAD(I,J)-PERM(I,J,ID))**2
133   CONTINUE
134   CONTINUE

      SR=0.0
      DO 136 I=1,16
      DO 135 J=1,4
         SR=SR+R(I,J)
135   CONTINUE
136   CONTINUE

      PMSR(S)=SQRT(SR/(16*4))

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******UNROTATED FACTOR PATTERNS, CALCULATE THE AVERAGE**********
*******EIGENVALUE RMSR FOR THE SAMPLE****************************

      PAMESR=0.0
      DO 222 J=1,SAMPLE
         PAMESR=PAMESR+PMESR(J)
222   CONTINUE

      PAMESR=PAMESR/SAMPLE

******CALCULATE THE STANDARD DEVIATION OF THE EIGENVALUE*********
******RMSR FOR THE SAMPLE OF OBSERVED UNROTATED PATTERNS*********

      USSE=0.0
      DO 335 J=1,SAMPLE
         USSE=USSE+(PMESR(J)-PAMESR)**2
335   CONTINUE

      SDUE=SQRT(USSE/(SAMPLE-1))

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******UNROTATED FACTOR PATTERNS, CALCULATE THE AVERAGE**********
*******LOADINGS RMSR FOR THE SAMPLE******************************

      PAMSR=0.0
      DO 955 J=1,SAMPLE
         PAMSR=PAMSR+PMSR(J)
955   CONTINUE

      PAMSR=PAMSR/SAMPLE

*****CALCULATE THE STANDARD DEVIATION OF THE LOADINGS***********
*****RMSR FOR THE SAMPLE OF OBSERVED UNROTATED PATTERNS*********

       PUSSM=0.0
       DO 953 J=1,SAMPLE
          PUSSM=PUSSM+(PMSR(J)-PAMSR)**2
953    CONTINUE

       PSDUM=SQRT(PUSSM/SAMPLE)

      END IF

******SET K TO 0 TO PERFORM RAW VARIMAX ROTATION ON EACH OF THE****
```

```
******OBSERVED SAMPLE UNROTATED FACTOR PATTERNS********************

      K=0
      CALL FROTA (NVAR, NF, A, LDA, NORM, MAXIT, W, EPS, B, LDB, T,
     &LDT)

******CALCULATE EIGENVALUES RAW VARIMAX ROTATED FACTOR PATTERN******
******THE MUST BE DONE TO CALCULATE MEASURES OF MERIT AND TO ENSURE*
******THE ROTATED FACTOR PATTERN IS IN THE CORRECT COLUMN ORDER*****
******SEE CHAPTER III FOR DETAILS

      DO 282 J=1,4
         NNUM(1,J,S)=0.0
      DO 267 I=1,16
         NNUM(1,J,S)=NNUM(1,J,S)+B(I,J)**2
267   CONTINUE
282   CONTINUE

******SORT EACH COLUMN OF THE EIGENVALUES OF THE ROTATED PATTERN****
******PUT FACTOR PATTERN COLUMN IN CORRESPONDING ORDER**************

       BIG=0.0

       DO 1121 J=1,NF
          IF (NNUM(1,J,S) .GT. BIG) THEN
       BIG=NNUM(1,J,S)
            AA=J
       END IF

1121   CONTINUE
       NNUM(1,AA,S)=0.0

       BIG=0.0

       DO 1123 J=1,NF
       IF (NNUM(1,J,S) .GT. BIG) THEN
          BIG=NNUM(1,J,S)
          BB=J
       END IF
1123   CONTINUE

       NNUM(1,BB,S)=0.0
       BIG=0.0
       DO 1125 J=1,NF
          IF (NNUM(1,J,S) .GT. BIG) THEN
          BIG=NNUM(1,J,S)
          CC=J
       END IF
125    CONTINUE

       NNUM(1,CC,S)=0.0
       BIG=0.0
       DO 1127 J=1,NF
       IF (NNUM(1,J,S) .GT. BIG) THEN
          BIG=NNUM(1,J,S)
          DD=J
       END IF
1127   CONTINUE

       NNUM(1,DD,S) =0.0

*******REDEFINE ROTATED FACTOR PATTERN IN CORRECT FORM**********

       DO 1128 I=1,NVAR
       CB(I,1)=B(I,AA)
       CB(I,2)=B(I,BB)
       CB(I,3)=B(I,CC)
       CB(I,4)=B(I,DD)
1128   CONTINUE

******WRITE CORRECTED RAW VARIMAX FACTOR PATTERN******************
```

```
*      PRINT*,'CORRECTED ROTATED RAW VARIMAX'
*      PRINT*
*      WRITE(*,10)((CB(I,J),J=1,NF),I=1,LDB )
*      PRINT*

******GENERATE PERMUTATIONS OF OUR SAMPLE RAW VARIMAX ROTATED*******
******FACTOR PATTERN IN PREPARATION FOR DETERMINING BEST PATTERN****

      DO 2211 I=1,NVAR
      DO 2210 J=1,NF
      PERM(I,J,1)=CB(I,J)
2210  CONTINUE
2211  CONTINUE

      DO 2212 I=1,16
      PERM(I,1,2)=CB(I,1)
      PFRM(I,2,2)=CB(I,2)
      PERM(I,3,2)=CB(I,4)
      PERM(I,4,2)=CB(I,3)
2212  CONTINUE

      DO 2213 I=1,16
      PERM(I,1,3)=CB(I,1)
      PERM(I,2,3)=CB(I,3)
      PERM(I,3,3)=CB(I,2)
      PERM(I,4,3)=CB(I,4)
2213  CONTINUE

      DO 2214 I=1,16
      PERM(I,1,4)=CB(I,1)
      PERM(I,2,4)=CB(I,3)
      PERM(I,3,4)=CB(I,4)
      PERM(I,4,4)=CB(I,2)
2214  CONTINUE

      DO 2215 I=1,16
      PERM(I,1,5)=CB(I,1)
      PERM(I,2,5)=CB(I,4)
      PERM(I,3,5)=CB(I,2)
      PERM(I,4,5)=CB(I,3)
2215  CONTINUE

      DO 2216 I=1,16
      PERM(I,1,6)=CB(I,1)
      PERM(I,2,6)=CB(I,4)
      PERM(I,3,6)=CB(I,3)
      PERM(I,4,6)=CB(I,2)
2216  CONTINUE

      DO 2217 I=1,16
      PERM(I,1,7)=CB(I,2)
      PERM(I,2,7)=CB(I,1)
      PERM(I,3,7)=CB(I,3)
      PERM(I,4,7)=CB(I,4)
2217  CONTINUE

      DO 2218 I=1,16
      PERM(I,1,8)=CB(I,2)
      PERM(I,2,8)=CB(I,1)
      PERM(I,3,8)=CB(I,4)
      PERM(I,4,8)=CB(I,3)
2218  CONTINUE

      DO 2219 I=1,16
      PERM(I,1,9)=CB(I,2)
      PERM(I,2,9)=CB(I,3)
      PERM(I,3,9)=CB(I,1)
      PERM(I,4,9)=CB(I,4)
2219  CONTINUE

      DO 2220 I=1,16
      PERM(I,1,10)=CB(I,2)
      PERM(I,2,10)=CB(I,3)
```

```
      PERM(I,3,10)=CB(I,4)
      PERM(I,4,10)=CB(I,1)
 2220 CONTINUE
      DO 2221 I=1,16
      PERM(I,1,11)=CB(I,2)
      PERM(I,2,11)=CB(I,4)
      PERM(I,3,11)=CB(I,1)
      PERM(I,4,11)=CB(I,3)
 2221 CONTINUE
      DO 2222 I=1,16
      PERM(I,1,12)=CB(I,2)
      PERM(I,2,12)=CB(I,4)
      PERM(I,3,12)=CB(I,3)
      PERM(I,4,12)=CB(I,1)
 2222 CONTINUE
      DO 2223 I=1,16
      PERM(I,1,13)=CB(I,3)
      PERM(I,2,13)=CB(I,1)
      PERM(I,3,13)=CB(I,2)
      PERM(I,4,13)=CB(I,4)
 2223 CONTINUE
      DO 2224 I=1,16
      PERM(I,1,14)=CB(I,3)
      PERM(I,2,14)=CB(I,1)
      PERM(I,3,14)=CB(I,4)
      PERM(I,4,14)=CB(I,2)
 2224 CONTINUE
      DO 2225 I=1,16
      PERM(I,1,15)=CB(I,3)
      PERM(I,2,15)=CB(I,2)
      PERM(I,3,15)=CB(I,1)
      PERM(I,4,15)=CB(I,4)
 2225 CONTINUE
      DO 2226 I=1,16
      PERM(I,1,16)=CB(I,3)
      PERM(I,2,16)=CB(I,2)
      PERM(I,3,16)=CB(I,4)
      PERM(I,4,16)=CB(I,1)
 2226 CONTINUE
      DO 2227 I=1,16
      PERM(I,1,17)=CB(I,3)
      PERM(I,2,17)=CB(I,4)
      PERM(I,3,17)=CB(I,1)
      PERM(I,4,17)=CB(I,2)
 2227 CONTINUE
      DO 2228 I=1,16
      PERM(I,1,18)=CB(I,3)
      PERM(I,2,18)=CB(I,4)
      PERM(I,3,18)=CB(I,2)
      PERM(I,4,18)=CB(I,1)
 2228 CONTINUE
      DO 2229 I=1,16
      PERM(I,1,19)=CB(I,4)
      PERM(I,2,19)=CB(I,1)
      PERM(I,3,19)=CB(I,2)
      PERM(I,4,19)=CB(I,3)
 2229 CONTINUE
      DO 2230 I=1,16
      PERM(I,1,20)=CB(I,4)
      PERM(I,2,20)=CB(I,1)
      PERM(I,3,20)=CB(I,3)
      PERM(I,4,20)=CB(I,2)
 2230 CONTINUE
```

```
      DO 2231 I=1,16
      PERM(I,1,21)=CB(I,4)
      PERM(I,2,21)=CB(I,2)
      PERM(I,3,21)=CB(I,1)
      PERM(I,4,21)=CB(I,3)
2231  CONTINUE

      DO 2232 I=1,16
      PERM(I,1,22)=CB(I,4)
      PERM(I,2,22)=CB(I,2)
      PERM(I,3,22)=CB(I,3)
      PERM(I,4,22)=CB(I,1)
2232  CONTINUE

      DO 2233 I=1,16
      PERM(I,1,23)=CB(I,4)
      PERM(I,2,23)=CB(I,3)
      PERM(I,3,23)=CB(I,1)
      PERM(I,4,23)=CB(I,2)
2233  CONTINUE

      DO 2234 I=1,16
      PERM(I,1,24)=CB(I,4)
      PERM(I,2,24)=CB(I,3)
      PERM(I,3,24)=CB(I,2)
      PERM(I,4,24)=CB(I,1)
2234  CONTINUE

******PUT ALL PERMUTTED MATRICES INTO BINARY FORM IN PREPARATION TO****
******POST MULTIPLY THE TRANSPOSE OF THE EXPERIMENTAL FACTOR PATTERN***
******TO DETERMINE BEST FACTOR PATTERN SCORE OF EACH OBSERVE RAW*******
******VARIMAX ROTATED FACTOR PATTERN***********************************

      DO 1053 H=1,24
      DO 1041 I=1,NVAR
      DO 1040 J=1,NF
         RV(I,J,H)=0
1040  CONTINUE
1041  CONTINUE

      DO 1052 I=1,NVAR
         BIG=ABS(PERM(I,1,H))
         LOCBIG=1
         RV(I,1,H)=1

      DO 1051 J=2,NF
      IF (ABS(PERM(I,J,H)) .GT. BIG) THEN
         RV(I,J,H)=1
         RV(I,LOCBIG,H)=0
         BIG=ABS(PERM(I,J,H))
         LOCBIG=J
      END IF
1051  CONTINUE
1052  CONTINUE
1053  CONTINUE

******CALCULATE PATTERN MATCHING SCORE FOR EACH PERMUTATION OF SAMPLE***
******OBSERVED RAW VARIMAX ROTATED FACTOR PATTERN***********************

      DO 3001 H=1,24
      DO 1152 I=1,4
         PRLOAD(I,1,H)=0
      DO 1151 J=1,NVAR
         PRLOAD(I,1,H)=PRLOAD(I,1,H)+(TRLOAD(I,J)*RV(J,1,H))
1151  CONTINUE
1152  CONTINUE

      DO 1154 I=1,4
         PRLOAD(I,2,H)=0
      DO 1153 J=1,NVAR
         PRLOAD(I,2,H)=PRLOAD(I,2,H)+(TRLOAD(I,J)*RV(J,2,H))
1153  CONTINUE
```

```
 1154  CONTINUE

       DO 1156 I=1,4
          PRLOAD(I,3,H)=0
       DO 1155 J=1,NVAR
          PRLOAD(I,3,H)=PRLOAD(I,3,H)+(TRLOAD(I,J)*RV(J,3,H))
 1155  CONTINUE
 1156  CONTINUE

       DO 1158 I=1,4
          PRLOAD(I,4,H)=0
       DO 1157 J=1,NVAR
          PRLOAD(I,4,H)=PRLOAD(I,4,H)+(TRLOAD(I,J)*RV(J,4,H))
 1157  CONTINUE
 1158  CONTINUE


       NUMHITS(H,S)=0
       DO 1159 I=1,4
       NUMHITS(H,S)=NUMHITS(H,S)+PRLOAD(I,I,H)
 1159  CONTINUE

 3001  CONTINUE

******DETERMINE BEST PERMUTATION BY DETERMINING HIGHEST SCORING******
******(NUMHITS) OF ALL PERMUTATIONS**********************************

       HRW=NUMHITS(1,S)
       ID=0
          DO 3600 H=2,24
       IF (NUMHITS(H,S) .GT. HRW) THEN
          HRW=NUMHITS(H,S)
          ID=H
       END IF
 3600  CONTINUE

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******UNROTATED FACTOR PATTERNS, CALCULATE THE AVERAGE**********
*******NUMBER OF MATCHES FOR THE SAMPLE**************************

       SUMPR=SUMPR+HRW
       IF (S .EQ. SAMPLE) THEN
          ABRHITS=FLOAT(SUMPR)/FLOAT(SAMPLE)
       END IF

******CALCULATE THE EIGENVALUES FOR THE BEST PERMUTED OBSERVED******
****** SAMPLE RAW VARIMAX ROTATED FACTOR PATTERN*********************

       DO 287 J=1,4
          PNNUM(1,J,S)=0.0
       DO 266 I=1,16
          PNNUM(1,J,S) =PNNUM(1,J,S)+PERM(I,J,ID)**2
 266   CONTINUE
 287   CONTINUE


******CALCULATE THE EIGENVALUE RMSR FOR EACH BEST PERMUTED SAMPLE**
******RAW VARIMAX ROTATED SAMPLE FACTOR PATTERN********************

       DO 492 J=1,4
          PRER(1,J,S)=(PNNUM(1,J,S)-INUM(J))**2
 492   CONTINUE

       PESR=0.0
       DO 493 J=1,4
          PESR=PESR+PRER(1,J,S)
 493   CONTINUE

       PRMESR(S)=SQRT(PESR/4)


******CALCULATE THE LOADING RMSR FOR EACH BEST PERMUTED SAMPLE*****
```

```
******RAW VARIMAX ROTATED SAMPLE FACTOR PATTERN********************

      DO 924 I=1,16
      DO 923 J=1,4
         R(I,J)=(LOAD(I,J)-PERM(I,J,ID))**2
923   CONTINUE
924   CONTINUE

      SR=0.0
      DO 926 I=1,16
      DO 925 J=1,4
         SR=SR+R(I,J)
925   CONTINUE
926   CONTINUE

      PRMSR(S)=SQRT(SR/(16*4))


*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE**********
*******RAW VARIMAX ROTATED FACTOR PATTERNS, CALCULATE THE AVERAGE*
*******EIGENVALUE RMSR FOR THE SAMPLE******************************

      IF (S .EQ. SAMPLE) THEN
         PRAMESR=0.0
      DO 888 J=1,SAMPLE
         PRAMESR=PRAMESR+PRMESR(J)
88C   CONTINUE

      PRAMESR=PRAMESR/SAMPLE

*****CALCULATE THE STANDARD DEVIATION OF THE LOADINGS*************
*****RMSR FOR THE SAMPLE OF OBSERVED RAW VARIMAX ROTATED PATTERNS*

      RSSE=0.0
      DO 1084 J=1,SAMPLE
         RSSE=RSSE+(PRMESR(J)-PRAMESR)**2
1084  CONTINUE

      SDRE=SQRT(RSSE/(SAMPLE-1))

      END IF

******SET K TO 1 TO PERFORM RAW VARIMAX ROTATION ON EACH OF THE****
******OBSERVED SAMPLE UNROTATED FACTOR PATTERNS********************

      K=1

      CALL FROTA (NVAR, NF, A, LDA, NORM, MAXIT, W, EPS, C, LDB, T,
     &LDT)

******CALCULATE EIGENVALUES NORMAL VARIMAX ROTATED FACTOR PATTERN***
******THE MUST BE DONE TO CALCULATE MEASURES OF MERIT AND TO ENSURE*
******THE ROTATED FACTOR PATTERN IS IN THE CORRECT COLUMN ORDER*****

      DO 283 J=1,4
         KNUM(1,J,S)=0.0
      DO 268 I=1,16
         KNUM(1,J,S) =KNUM(1,J,S)+C(I,J)**2
268   CONTINUE
283   CONTINUE

******SORT EACH COLUMN OF THE EIGENVALUES OF THE ROTATED PATTERN****
******PUT FACTOR PATTERN COLUMN IN CORRESPONDING ORDER**************

       BIG=0.0

       DO 1221 J=1,NF
       IF (KNUM(1,J,S) .GT. BIG) THEN
          BIG=KNUM(1,J,S)
          AA=J
       END IF

1221   CONTINUE
```

```
      KNUM(1,AA,S)=0.0

      BIG=0.0
      DO 1223 J=1,NF
      IF (KNUM(1,J,S) .GT. BIG) THEN
         BIG=KNUM(1,J,S)
         BB=J
      END IF
1223  CONTINUE

      KNUM(1,BB,S)=0.0
      BIG=0.0
      DO 1225 J=1,NF
      IF (KNUM(1,J,S) .GT. BIG) THEN
         BIG=KNUM(1,J,S)
         CC=J
      END IF
1225  CONTINUE

      KNUM(1,CC,S)=0.0
      BIG=0.0
      DO 1227 J=1,NF
      IF (KNUM(1,J,S) .GT. BIG) THEN
         BIG=KNUM(1,J,S)
         DD=J
      END IF
1227  CONTINUE

      KNUM(1,DD,S) =0.0

      DO 1228 I=1,NVAR
      CK(I,1)=C(I,AA)
      CK(I,2)=C(I,BB)
      CK(I,3)=C(I,CC)
      CK(I,4)=C(I,DD)
1228  CONTINUE

******WRITE CORRECTED NORMAL VARIMAX FACTOR PATTERN*******************

*     PRINT*,'CORRECTED ROTATED NORMAL VARIMAX'
*     PRINT*
*     WRITE(*,10)((CK(I,J),J=1,NF),I=1,LDB )
*     PRINT*

******WRITE UNCORRECTED NORMAL VARIMAX FACTOR PATTERN*******************

*     PRINT*,'UNCORRECTED ROTATED NORMAL VARIMAX'
*     PRINT*
*     WRITE(*,10)((C(I,J),J=1,NF),I=1,LDB )
*     PRINT*

******CALCULATE EACH PERMUTATION OF NORMAL VARIMAX ROTATED************
******FACTOR PATTERN**************************************************

      DO 2111 I=1,NVAR
      DO 2110 J=1,NF
      PERM(I,J,1)=CK(I,J)
2110  CONTINUE
2111  CONTINUE

      DO 2112 I=1,16
      PERM(I,1,2)=CK(I,1)
      PERM(I,2,2)=CK(I,2)
      PERM(I,3,2)=CK(I,4)
      PERM(I,4,2)=CK(I,3)
2112  CONTINUE

      DO 2113 I=1,16
      PERM(I,1,3)=CK(I,1)
      PERM(I,2,3)=CK(I,3)
      PERM(I,3,3)=CK(I,2)
```

```
      PERM(I,4,3)=CK(I,4)
 2113 CONTINUE

      DO 2114 I=1,16
      PERM(I,1,4)=CK(I,1)
      PERM(I,2,4)=CK(I,3)
      PERM(I,3,4)=CK(I,4)
      PERM(I,4,4)=CK(I,2)
 2114 CONTINUE

      DO 2115 I=1,16
      PERM(I,1,5)=CK(I,1)
      PERM(I,2,5)=CK(I,4)
      PERM(I,3,5)=CK(I,2)
      PERM(I,4,5)=CK(I,3)
 2115 CONTINUE

      DO 2116 I=1,16
      PERM(I,1,6)=CK(I,1)
      PERM(I,2,6)=CK(I,4)
      PERM(I,3,6)=CK(I,3)
      PERM(I,4,6)=CK(I,2)
 2116 CONTINUE

      DO 2117 I=1,16
      PERM(I,1,7)=CK(I,2)
      PERM(I,2,7)=CK(I,1)
      PERM(I,3,7)=CK(I,3)
      PERM(I,4,7)=CK(I,4)
 2117 CONTINUE

      DO 2118 I=1,16
      PERM(I,1,8)=CK(I,2)
      PERM(I,2,8)=CK(I,1)
      PERM(I,3,8)=CK(I,4)
      PERM(I,4,8)=CK(I,3)
 2118 CONTINUE

      DO 2119 I=1,16
      PERM(I,1,9)=CK(I,2)
      PERM(I,2,9)=CK(I,3)
      PERM(I,3,9)=CK(I,1)
      PERM(I,4,9)=CK(I,4)
 2119 CONTINUE

      DO 2120 I=1,16
      PERM(I,1,10)=CK(I,2)
      PERM(I,2,10)=CK(I,3)
      PERM(I,3,10)=CK(I,4)
      PERM(I,4,10)=CK(I,1)
 2120 CONTINUE

      DO 2121 I=1,16
      PERM(I,1,11)=CK(I,2)
      PERM(I,2,11)=CK(I,4)
      PERM(I,3,11)=CK(I,1)
      PERM(I,4,11)=CK(I,3)
 2121 CONTINUE

      DO 2122 I=1,16
      PERM(I,1,12)=CK(I,2)
      PERM(I,2,12)=CK(I,4)
      PERM(I,3,12)=CK(I,3)
      PERM(I,4,12)=CK(I,1)
 2122 CONTINUE

      DO 2123 I=1,16
      PERM(I,1,13)=CK(I,3)
      PERM(I,2,13)=CK(I,1)
      PERM(I,3,13)=CK(I,2)
      PERM(I,4,13)=CK(I,4)
 2123 CONTINUE
```

```
      DO 2124 I=1,16
      PERM(I,1,14)=CK(I,3)
      PERM(I,2,14)=CK(I,1)
      PERM(I,3,14)=CK(I,4)
      PERM(I,4,14)=CK(I,2)
 2124 CONTINUE
      DO 2125 I=1,16
      PERM(I,1,15)=CK(I,3)
      PERM(I,2,15)=CK(I,2)
      PERM(I,3,15)=CK(I,1)
      PERM(I,4,15)=CK(I,4)
 2125 CONTINUE
      DO 2126 I=1,16
      PERM(I,1,16)=CK(I,3)
      PERM(I,2,16)=CK(I,2)
      PERM(I,3,16)=CK(I,4)
      PERM(I,4,16)=CK(I,1)
 2126 CONTINUE
      DO 2127 I=1,16
      PERM(I,1,17)=CK(I,3)
      PERM(I,2,17)=CK(I,4)
      PERM(I,3,17)=CK(I,1)
      PERM(I,4,17)=CK(I,2)
 2127 CONTINUE
      DO 2128 I=1,16
      PERM(I,1,18)=CK(I,3)
      PERM(I,2,18)=CK(I,4)
      PERM(I,3,18)=CK(I,2)
      PERM(I,4,18)=CK(I,1)
 2128 CONTINUE
      DO 2129 I=1,16
      PERM(I,1.19)=CK(I,4)
      PERM(I,2,19)=CK(I,1)
      PERM(I,3,19)=CK(I,2)
      PERM(I,4,19)=CK(I,3)
 2129 CONTINUE
      DO 2130 I=1,16
      PERM(I,1,20)=CK(I,4)
      PERM(I,2,20)=CK(I,1)
      PERM(I,3,20)=CK(I,3)
      PERM(I,4,20)=CK(I,2)
 2130 CONTINUE
      DO 2131 I=1,16
      PERM(I,1,21)=CK(I,4)
      PERM(I,2,21)=CK(I,2)
      PERM(I,3,21)=CK(I,1)
      PERM(I,4,21)=CK(I,3)
 2131 CONTINUE
      DO 2132 I=1,16
      PERM(I,1,22)=CK(I,4)
      PERM(I,2,22)=CK(I,2)
      PERM(I,3,22)=CK(I,3)
      PERM(I,4,22)=CK(I,1)
 2132 CONTINUE
      DO 2133 I=1,16
      PERM(I,1,23)=CK(I,4)
      PERM(I,2,23)=CK(I,3)
      PERM(I,3,23)=CK(I,1)
      PERM(I,4,23)=CK(I,2)
 2133 CONTINUE
      DO 2134 I=1,16
      PERM(I,1,24)=CK(I,4)
      PERM(I,2,24)=CK(I,3)
```

```
      PERM(I,3,24)=CK(I,2)
      PERM(I,4,24)=CK(I,1)
2134  CONTINUE

******PUT ALL PERMUTTED MATRICES INTO BINARY FORM IN PREPARATION TO****
******POST MULTIPLY THE TRANSPOSE OF THE EXPERIMENTAL FACTOR PATTERN***
******TO DETERMINE BEST FACTOR PATTERN SCORE***************************

      DO 1953 H=1,24
      DO 1941 I=1,NVAR
      DO 1940 J=1,NF
         KV(I,J,H)=0
1940  CONTINUE
1941  CONTINUE

      DO 1952 I=1,NVAR
         BIG=ABS(PERM(I,1,H))
         LOCBIG=1
         KV(I,1,H)=1
      DO 1951 J=2,NF
      IF (ABS(PERM(I,J,H)) .GT. BIG) THEN
         KV(I,J,H)=1
         KV(I,LOCBIG,H)=0
         BIG=ABS(PERM(I,J,H))
         LOCBIG=J
      END IF
1951  CONTINUE
1952  CONTINUE
1953  CONTINUE

******CALCULATE PATTERN MATCHING SCORE FOR EACH PERMUTATION OF SAMPLE***
******OBSERVED NORMAL VARIMAX ROTATED FACTOR PATTERN********************

      DO 3002 H=1,24
      DO 1162 I=1,4
         PRLOAD(I,1,H)=0
      DO 1161 J=1,NVAR
         PRLOAD(I,1,H)=PRLOAD(I,1,H)+(TRLOAD(I,J)*KV(J,1,H))
1161  CONTINUE
1162  CONTINUE

      DO 1164 I=1,4
         PRLOAD(I,2,H)=0
      DO 1163 J=1,NVAR
         PRLOAD(I,2,H)=PRLOAD(I,2,H)+(TRLOAD(I,J)*KV(J,2,H))
1163  CONTINUE
1164  CONTINUE

      DO 1166 I=1,4
         PRLOAD(I,3,H)=0
      DO 1165 J=1,NVAR
         PRLOAD(I,3,H)=PRLOAD(I,3,H)+(TRLOAD(I,J)*KV(J,3,H))
1165  CONTINUE
1166  CONTINUE

      DO 1168 I=1,4
         PRLOAD(I,4,H)=0
      DO 1167 J=1,NVAR
         PRLOAD(I,4,H)=PRLOAD(I,4,H)+(TRLOAD(I,J)*KV(J,4,H))
1167  CONTINUE
1168  CONTINUE

      NUMHITS(H,S)=0
      DO 1169 I=1,4
      NUMHITS(H,S)=NUMHITS(H,S)+PRLOAD(I,I,H)
1169  CONTINUE
3002  CONTINUE

******DETERMINE BEST PERMUTATION BY DETERMINING HIGHEST SCORING*****
******(NUMHITS) OF ALL PERMUTATIONS*********************************

      HKI=NUMHITS(1,S)
      ID=0
```

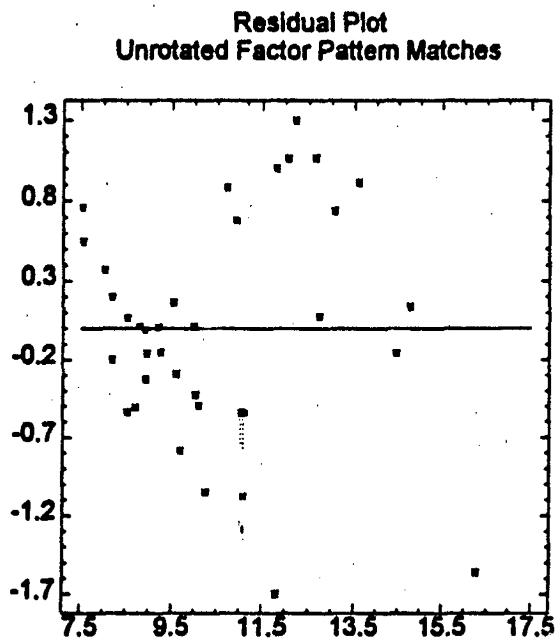
```
      DO 3700 H=2,24
         IF (NUMHITS(H,S) .GT. HKI) THEN
            HKI=NUMHITS(H,S)
            ID=H
         END IF
3700  CONTINUE

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******NORMAL VARIMAX ROTATED FACTOR PATTERNS, CALCULATE ********
*******THE AVERAGE NUMBER OF MATCHES FOR THE SAMPLE**************

      SUMPK=SUMPK+HKI
      IF (S .EQ. SAMPLE) THEN
         ABKHITS=FLOAT(SUMPK)/FLOAT(SAMPLE)
      END IF

*******CALCULATE EIGENVALUES BEST PERMUTTED KAISER VARIMAX*************

      DO 983 J=1,4
         PKNUM(1,J,S)=0.0
      DO 968 I=1,16
         PKNUM(1,J,S) =PKNUM(1,J,S)+PERM(I,J,ID)**2
968   CONTINUE
983   CONTINUE

******CALCULATE THE EIGENVALUE RMSR FOR EACH BEST PERMUTED SAMPLE**
******NORMAL VARIMAX ROTATED SAMPLE FACTOR PATTERN*****************

      DO 442 J=1,4
         PRER(1,J,S)=(PKNUM(1,J,S)-INUM(J))**2
442   CONTINUE

      PESR=0.0
         DO 443 J=1,4
      PESR=PESR+PPER(1,J,S)
443   CONTINUE

      PKMESR(S)=SQRT(PESR/4)

******CALCULATE THE LOADING RMSR FOR EACH BEST PERMUTED SAMPLE*****
******NORMAL VARIMAX ROTATED SAMPLE FACTOR PATTERN*****************

      DO 914 I=1,16
      DO 913 J=1,4
         R(I,J)=(LOAD(I,J)-PERM(I,J,ID))**2
913   CONTINUE
914   CONTINUE

      SR=0.0
      DO 916 I=1,16
      DO 915 J=1,4
         SR=SR+R(I,J)
915   CONTINUE
916   CONTINUE

      PKMSR(S)=SQRT(SR/(16*4))

********IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******NORMAL VARIMAX ROTATED FACTOR PATTERNS, CALCULATE THE*****
*******AVERAGE EIGENVALUE RMSR FOR THE SAMPLE********************

      IF (S .EQ. SAMPLE) THEN

      PKAMESR=0.0
      DO 445 J=1,SAMPLE
         PKAMESR=PKAMESR+PKMESR(J)
445   CONTINUE

      PKAMESR=PKAMESR/SAMPLE

******CALCULATE THE STANDARD DEVIATION OF THE EIGENVALUE*********
```

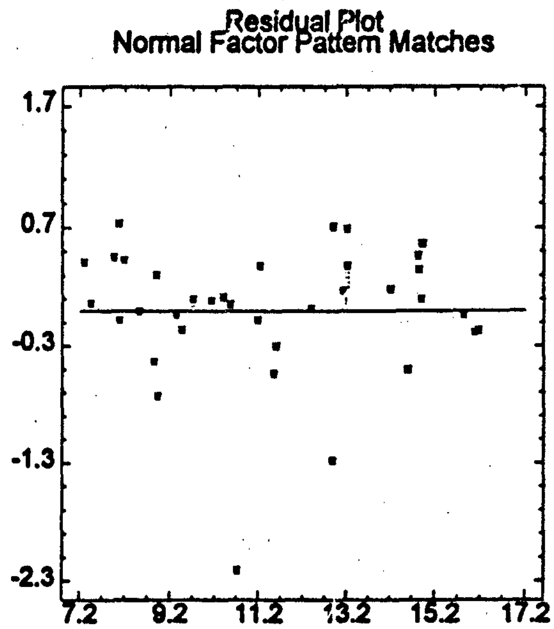
```
******RMSR FOR THE SAMPLE OF OBSERVED UNROTATED PATTERNS*********

      KSSE=0.0
      DO 664 J=1,SAMPLE
         KSSE=KSSE+(PKMESR(J)-PKAMESR)**2
664   CONTINUE

      SDKE=SQRT(KSSE/(SAMPLE-1))

*******IF WE ARE NOW WORKING WITH THE LAST OF OUR SAMPLE*********
*******NORMAL VARIMAX FACTOR PATTERNS, CALCULATE THE AVERAGE*****
*******LOADINGS RMSR FOR THE SAMPLE******************************

      PKAMSR=0.0
      DO 956 J=1,SAMPLE
         PKAMSR=PKAMSR+PKMSR(J)
956   CONTINUE
      PKAMSR=PKAMSR/SAMPLE

*****CALCULATE THE STANDARD DEVIATION OF THE LOADINGS***********
*****RMSR FOR THE SAMPLE OF OBSERVED UNROTATED PATTERNS*********

      PKSSM=0.0
      DO 973 J=1,SAMPLE
         PKSSM=PKSSM+(PKMSR(J)-PKAMSR)**2
973   CONTINUE
      PSDKM=SQRT(PKSSM/(SAMPLE-1))

      END IF

******END LOOP FOR READING 100 SAMPLES********************

300   CONTINUE
      CLOSE(2)

******WRITE DATA TO 'out.dat' FILE************************

******COUNT=STRUCTUR NUMBER

******HINDEX=UNIQUENESS INDEX

******CINDEX=COMPLEXITY INDEX

******AIBHITS=AVERAGE UNROTATED MATCHING SCORE

******PAMSR=AVERAGE UNROTATED LOADINGS RMSR

******PSDUM=STD DEV UNROTATED LOADINGS RMSR

******PAMESR=AVERAGE UNROTATED EIGENVALUE RMSR

******SDUE=STD DEV UNROTATED EIGENVALUE RMSR

******ABRHITS=AVERAGE ROTATED RAW MATCHING SCORE

******PRAMSR=AVERAGE ROTATED RAW LOADINGS RMSR

******PSDRM=STD DEV ROTATED RAW LOADINGS RMSR

******PRAMESR=AVERAGE ROTATED RAW EIGENVALUE RMSR

******SDRE=STD DEV ROTATED RAW EIGENVALUE RMSR

******ABKHITS=AVERAGE ROTATED NORMAL MATCHING SCORE

******PKAMSR=AVERAGE ROTATED NORMAL LOADINGS RMSR

******PSDKM=STD DEV ROTATED NORMAL LOADINGS RMSR

******PKAMESR=AVERAGE ROTATED NORMAL EIGENVALUE RMSR

******SDKE=STD DEV ROTATED NORMAL EIGENVALUE RMSR
```

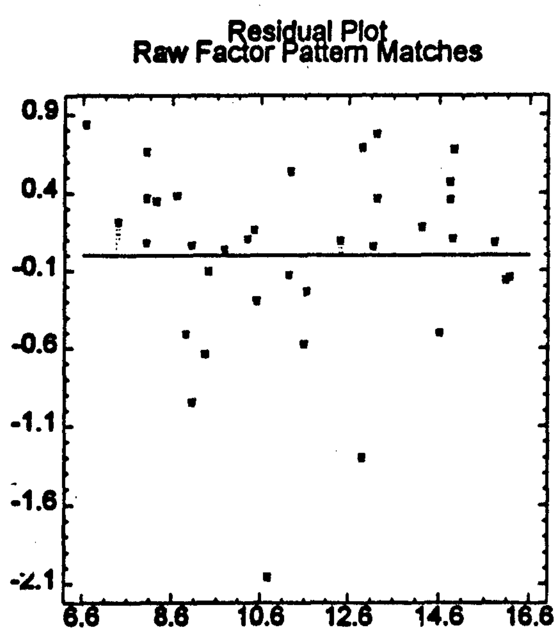
```
      WRITE(26,7006)COUNT,STRAT,HINDEX,CINDEX,AIBHITS,
     &PAMSR,PSDUM,PAMESR,SDUE,
     &ABRHITS,PRAMSR,PSDRM,PRAMESR,
     &SDRE,ABKHITS,PKAMSR,PSDKM,PKAMESR,SDKE

7006  FORMAT(1XI3,1XF6.3,1XF6.3,1XF6.3,1XF5.1,
     &1XF9.5,1XF9.5,1XF9.5,1XF9.5,1XF5.1,1XF9.5,1XF9.5,
     &1XF9.5,1XF9.5,1XF5.1,1XF9.5,1XF9.5,1XF9.5,1XF9.5)

      CLOSE(26)
      END
```

Appendix E. Residual Plots

This Appendix illustrates the residual plots for the pattern matching regression model.

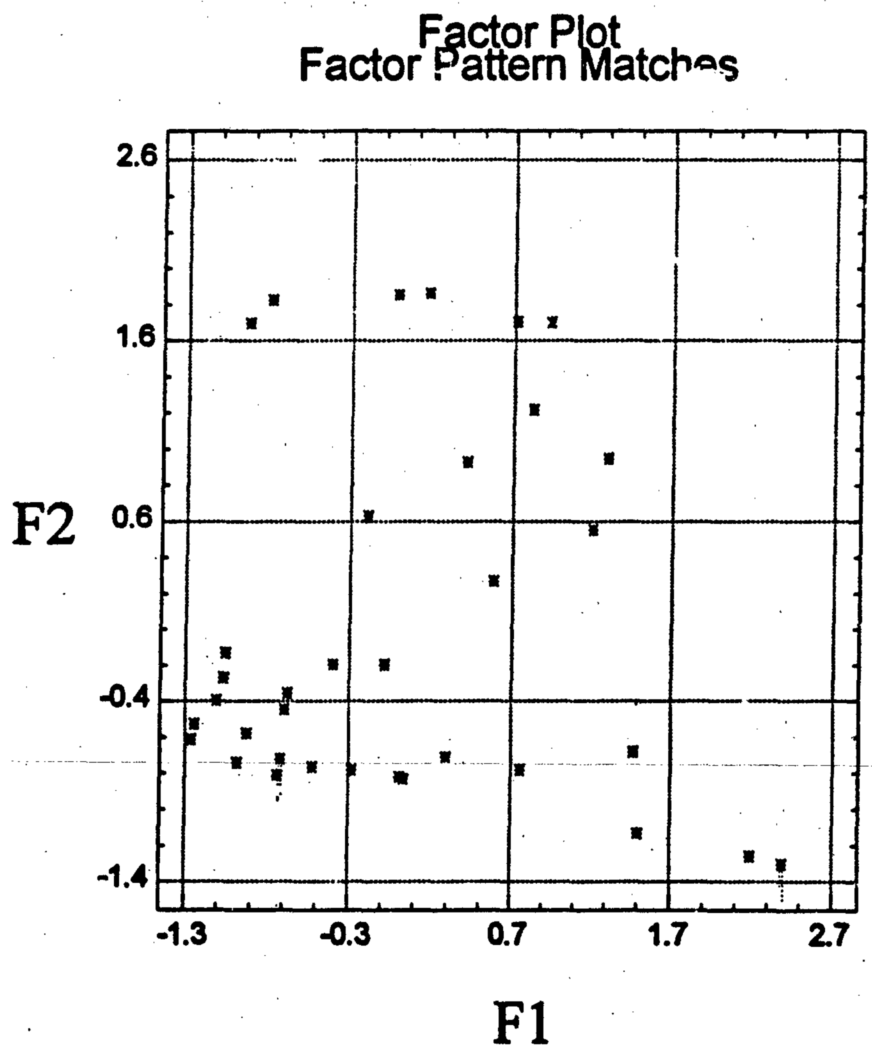






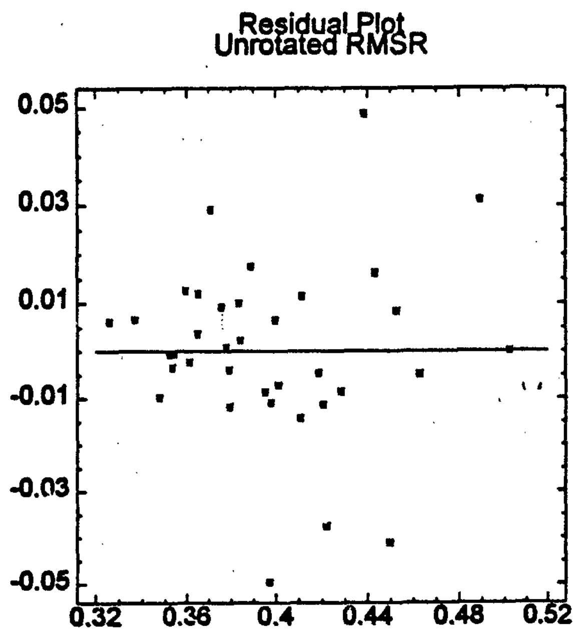
Appendix F. Factor Plots

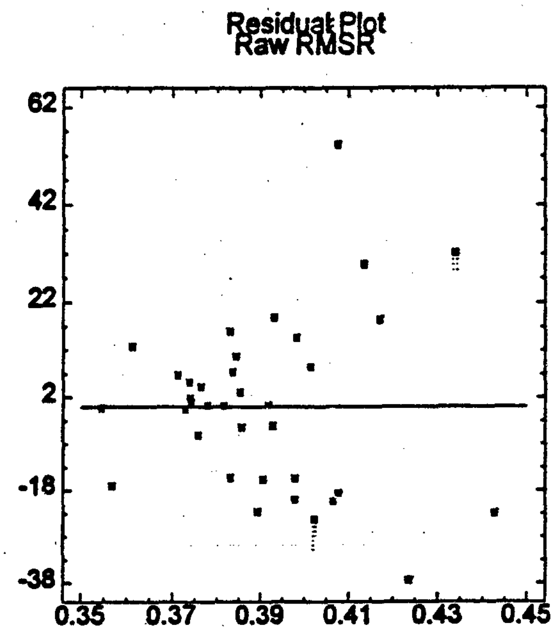
This Appendix illustrates the Factor plot for the pattern matching score factor Analysis.

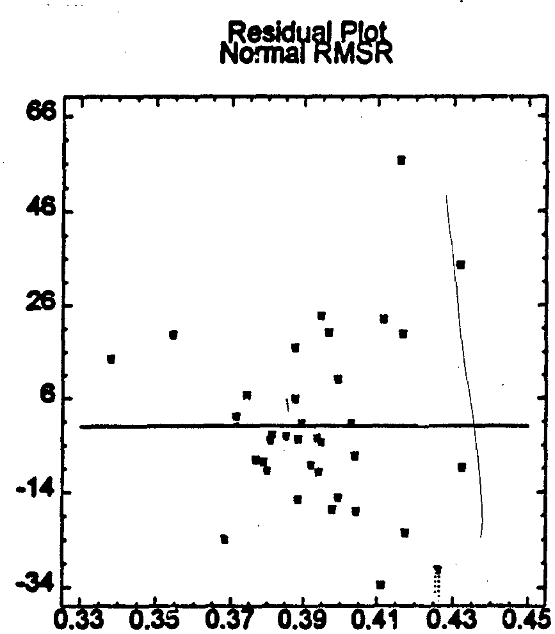


Appendix G. Residual Plots

This Appendix illustrates the residual plots for the RMSR regression model.

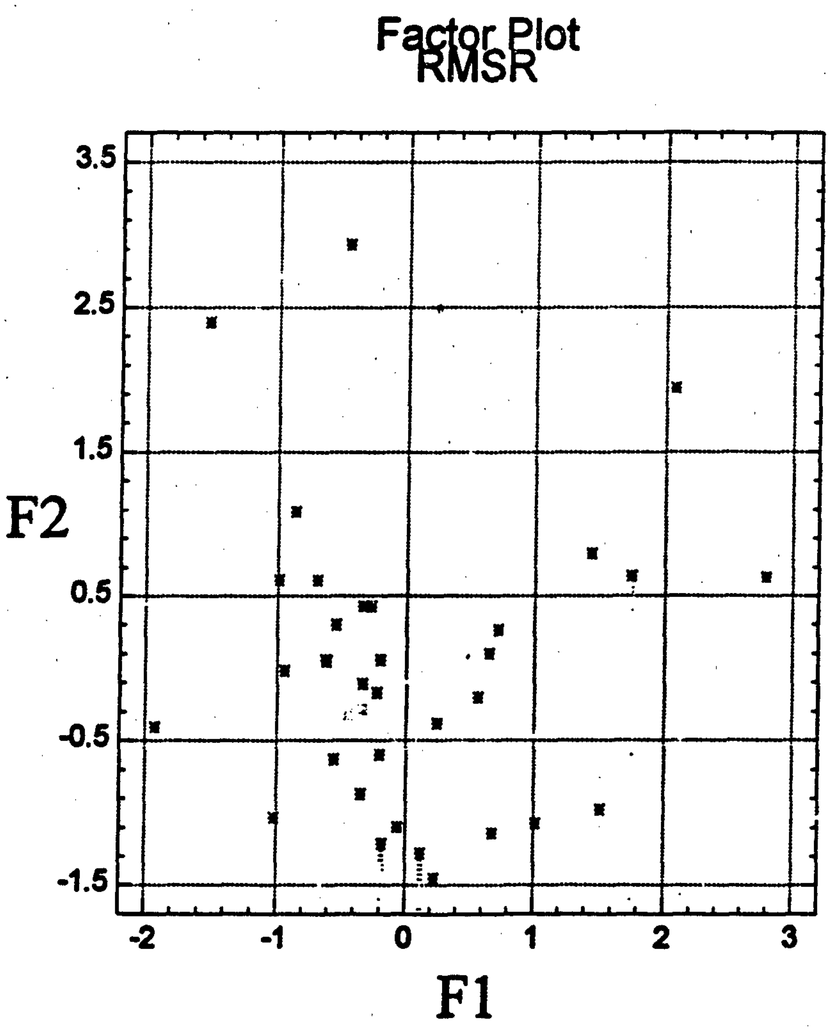






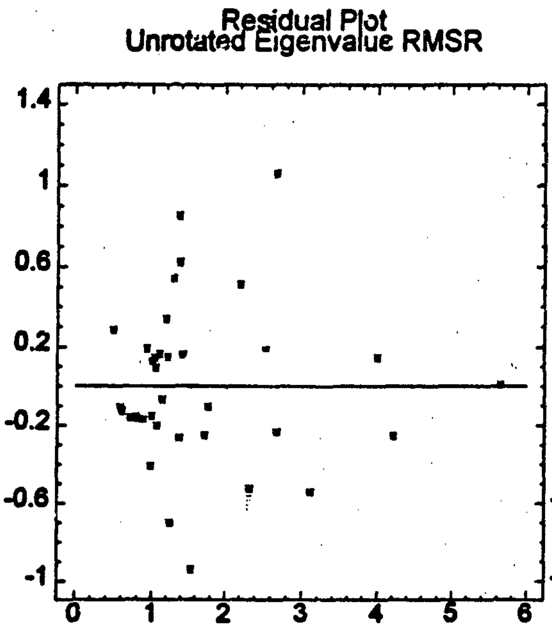
Appendix H. Factor Plots

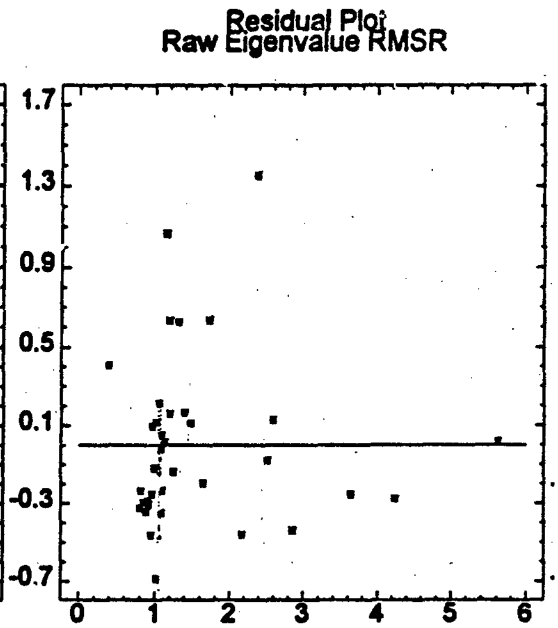
This Appendix illustrates the Factor plot for the RMSR factor Analysis.

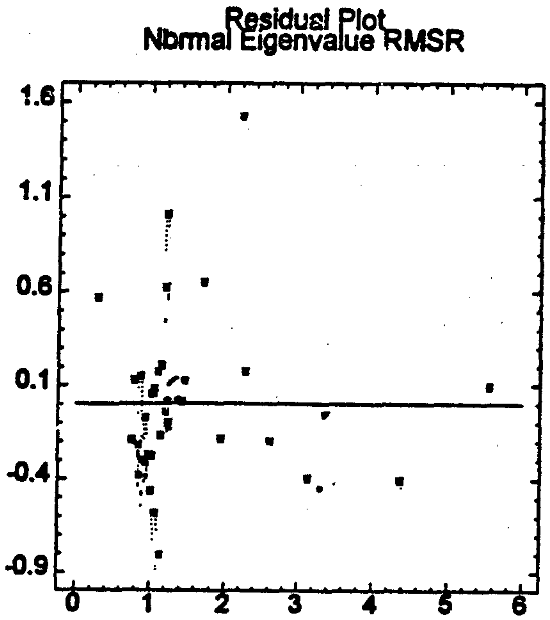


Appendix I. Residual Plots

This Appendix illustrates the residual plots for the RMSER regression model.

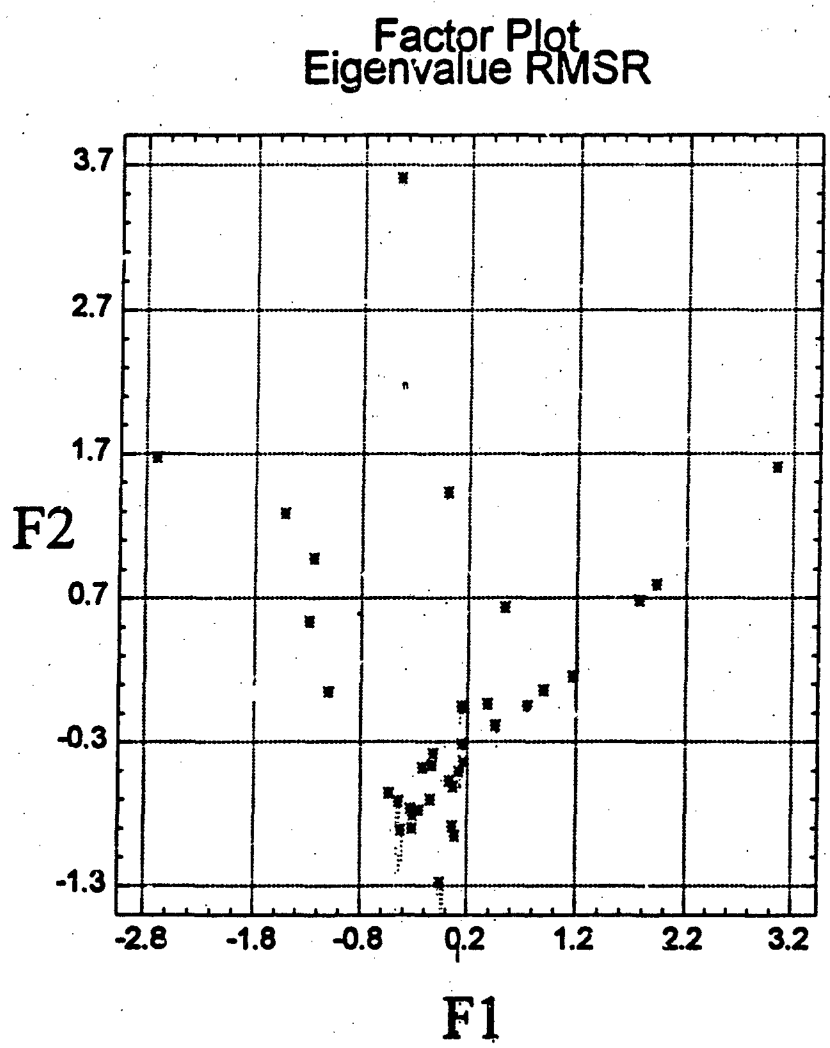






Appendix J. Factor Plots

This Appendix illustrates the Factor plot for the RMSER factor Analysis.



Appendix K. Kaiser Data

K.1 Correlation Matrix

This Appendix lists the correlation matrix for the Holzinger and Harman psychological tests (Holzinger and Harman, 1941). Note the matrix has been split into two (24*8) matrices to fit on the page.

1	.318	.403	.468	.321	.335	.304	.332	.326	.116	.308	.314
.318	1	.317	.230	.285	.234	.157	.157	.195	.057	.150	.145
.403	.317	1	.305	.247	.268	.223	.382	.184	-.075	.091	.140
.468	.230	.305	1	.227	.327	.335	.391	.325	.099	.110	.160
.321	.285	.247	.227	1	.622	.656	.578	.723	.311	.344	.215
.335	.234	.268	.327	.622	1	.722	.527	.714	.203	.353	.095
.304	.157	.223	.335	.656	.722	1	.619	.685	.246	.232	.181
.332	.157	.382	.391	.578	.527	.619	1	.532	.285	.300	.271
.326	.195	.184	.325	.723	.714	.685	.532	1	.170	.280	.113
.116	.057	-.075	.099	.311	.203	.246	.285	.170	1	.484	.585
.308	.150	.091	.110	.344	.353	.232	.300	.280	.484	1	.428
.314	.145	.140	.160	.215	.095	.181	.271	.113	.585	.428	1
.489	.239	.321	.327	.344	.309	.345	.395	.280	.403	.535	.512
.125	.103	.177	.066	.280	.292	.236	.252	.260	.172	.350	.131
.238	.131	.065	.127	.229	.251	.172	.175	.248	.154	.240	.173
.414	.272	.263	.322	.187	.291	.180	.296	.242	.124	.314	.119
.176	.005	.177	.187	.208	.273	.228	.255	.274	.289	362	.276
.368	.255	.211	.251	.263	.167	.159	.250	.208	.317	.350	.349
.270	.112	.312	.137	.190	.251	.226	.274	.274	.190	.290	.110
.365	.292	.297	.339	.398	.435	.451	.427	.446	.173	.202	.246
.369	.306	.165	.349	.318	.263	.314	.362	.266	.405	.399	.355
.413	.232	.250	.380	.441	.386	.396	.357	.483	.160	.304	.193
.474	.348	.383	.335	.435	.431	.405	.501	.504	.262	.251	.350
.282	.211	.203	.248	.420	.433	.437	.388	.424	.531	.412	.414

439	.125	.238	.414	.176	.368	.270	.365	.369	.413	.474	.282
.239	.103	.131	.272	.005	.255	.112	.292	.306	.232	.348	.211
.321	.177	.065	.263	.177	.211	.312	.297	.165	.250	.383	.203
.327	.066	.127	.322	.187	.251	.137	.339	.349	.380	.335	.248
.344	.280	.229	.187	.208	.263	.190	.398	.318	.441	.435	.420
.309	.292	.251	.291	.273	.167	.251	.435	.263	.386	.431	.433
.345	.236	.172	.180	.228	.159	.226	.451	.314	.396	.405	.437
.395	.252	.175	.296	.255	.250	.274	.427	.362	.357	.501	.388
.280	.260	.248	.242	.274	.208	.274	.446	.266	.483	.504	.424
.408	.172	.154	.124	.289	.317	.190	.173	.405	.160	.262	.531
.535	.350	.240	.314	.362	.350	.290	.202	.399	.304	.251	.412
.512	.131	.173	.119	.278	.349	.110	.246	.355	.193	.350	.414
1	.195	.139	.281	.194	.323	.263	.241	.425	.279	.382	.358
.195	1	.370	.412	.341	.201	.200	.302	.183	.243	.242	.304
.139	.370	1	.325	.345	.334	.192	.272	.232	.246	.256	.165
.281	.412	.325	1	.324	.344	.258	.388	.348	.283	.360	.262
.194	.341	.345	.324	1	.448	.324	.262	.173	.273	.287	.326
.323	.201	.334	.344	.448	1	.358	.301	.357	.317	.272	.405
.263	.200	.192	.258	.324	.358	1	.167	.331	.342	.303	.374
.241	.302	.272	.388	.262	.301	.167	1	.413	.463	.509	.366
.425	.183	.232	.348	.173	.357	.331	.413	1	.374	.451	.448
.279	.243	.246	.283	.273	.317	.342	.463	.374	1	.503	.375
.382	.242	.256	.360	.287	.272	.303	.509	.451	.503	1	.434
.358	.304	.165	.262	.326	.405	.374	.366	.448	.375	.434	1

Bibliography

1. Bauer, Lt. Col. Keneth W. Lecture Notes, OPER 685, Applied Mutivariate Analysis. School of Engineering, Air Force Institute of Technology, Wright Patterson AFB OH, Winter Quarter, 1993.

2. Bauer, Capt Keneth W. A Monte Carlo Study of Dimensionality Assessment and Factor Interpretation in Factor Analysis. MS thesis, AFIT/GOR/OS/81D-9. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, December 1981.

3. Burden, Richard L. and J. Douglas Faires. Numerical Analysis. 4th ed. Boston: PWS-KENT Publishing Co., 1989.

4. Carroll, J.B. "An Analytical Solution for Approximating Simple Structure in Factor Analysis," Psychometrika, Volume 18, 23-38, (March,1953).

5. Catell, R.B. "The Scree Test for the Number of Factors," Multivariate Behavioral Research, 245-276, (April 1966.

6. Cliff, Norman and Roger Pennell, "The Influence of Communality, Factor Strength, and Loading size on the Sampling Characteristics of Factor Loadings," Psychometrika, Volume 32, 309-326 (1967).

7. Dillon, William R. and Matthew Goldstein. Multivariate Analysis Methods and Applications. New York: John Wiley & Sons, 1972.

8. Fidell, Linda S. and Barbara G. Tabachnick. Using Multivariate Statistics. New York: Harper & Row, Publishers, 1983.

9. Fruchter, Benjamin. Introduction to Factor Analysis. New York: D. Van Nostrand Company, Inc., 1954.

10. Holtzinger, K.J. and H.H. Harman. Factor Analysis. Chicago: Univ. Chicago Press, 1941.

11. Horn, John L. "A Rationale and Test for the Number of Factors in Factor analysis," Psychometrika, Volume 30, 179-185, (June 1966).

12. IMSL, "Fortran Routines for Mathematical Applications,"STAT/LIBRARY, Volume 3, 1987.

13. Jackson, Edward J. A User's Guide to Principal Components. New York: John Wiley & Sons, 1991.

14. Kaiser, H.F. "The Varimax Criterion for Analytic Rotation in Factor Analysis," Psychometrika, Volume23, 187-200, (September 1958).

15. Kaiser, H.F. "The Application of Electronic Computers to Factor Analysis," Educational and Educational and Psychological Measurement, Volume 20, 141-151, (1960).

16. Kelton, Averill M. and W. David Kelton. Simulation Modeling and Analysis. 2nd ed. McGraw-Hill Inc., 1991.

17. Lawley, D.N. and A.E. Maxwell. Factor Analysis as a Statistical Method 2nd ed. New York: American Elsevier Publishing Co., 1971.

18. MacCallum, R.C. and Cornelius, E.T.A., III "A Monte Carlo Investigation of Recovery of Structure by ALSCAL," Psychometrika, Volume 42, 401-428, (1977).

19. Morrison, Donald F. Multivariate Statistical Methods. 3rd ed. New York: McGraw-Hill Publishing Co., 1990.

20. Odaka Y. and Yutaka Tanaka. "Influential Observations in Principal Factor Analysis," Psychometrika, Volume 54, 475-485 (September 1989).

21. Pennell, Roger "The Influence of Communality and N on the Sampling Distributions of Factor Loadings," Psychometrika, Volume 33, 423-439, (1968).

22. Saunders, David R. "A Computer Program to Find the Best-Fitting Orthogonal Factors for a Given Hypothesis," Psychometrika, Volume25, 19[illegible]05, (June 1960).

23. Tucker, Ledyard R. and others. "Evaluation of Factor Analytic Research Procedures by Means of Simulated Correlation Matrices," Psychometrika, Volume 25, 199-205 (June 1960).

Vita

Captain William M. Ibinson was born on 11 October, 1960 in Pittsburgh, Pennsylvania. He graduated from Richland Senior High School in 1978 and enlisted in the United States Air Force. After completing Basic Military Training, he attended the Aircraft Armament Systems School, Lowry AFB, Colorado where he was a distinguished graduate. Captain Ibinson's first assignment was to the 5th Fighter Intercept Squadron at Minot AFB, North Dakota as an armament specialist for the Convair F-106 Delta Dart. In 1982 he received an active duty appointment to the United States Air Force Academy, graduating with a Master of Sciences in Management Sciences.

As a newly commissioned officer, Captain Ibinson was then assigned to the Technology Assessment Division, Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio where he conducted mission analysis in support of future airlift requirements for Special Operations Forces. This was followed by a tour at the Aircraft Survivability and Research Facility at Wright-Patterson AFB. As director of facility operations, Captain Ibinson supervised Air Force, government and contract personnel conducting Joint Live-Fire survivability testing For the C-17, F-15, F-16, and AMRAAM weapon systems.

In 1991 Captain Ibinson entered the School of Engineering in Operations Research, Air Force Institute of Technology.

Permanent address: 3734 N. Woodland Cir.
Gibsonia, PA 15044

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1993

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
A MONTE CARLO STUDY OF THE RAW AND NORMAL VARIMAX ROTATION CRITERION IN FACTOR ANALYSIS

5. FUNDING NUMBERS

6. AUTHOR(S)
William M. Ibinson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GOR/ENS/93M-27

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Factor analysis is a multivariate statistical procedure for analyzing and reducing large data sets. Many factor analysis schemes and techniques are available that lead to strikingly different results from the same data. This research effort used a Monte Carlo approach to investigate the properties of two rotation methods for simple structure, Kaiser's raw and normal varimax criterion. Data sets were developed from a set of contrived experimental factor patterns by multiplying each factor pattern by its transpose to create a covariance matrix. Data sets of multivariate normal deviates were in turn generated from each covariance matrix via the Choleski algorithm. Rotated factor pattern matrices from each data set were compared to their respective experimental factor pattern on the basis of structure, loadings and eigenvalues. These performance issues are addressed through regression analysis and separate factor analysis in which the grand mean of proposed measures of effectiveness are predicted. These measures of effectiveness include structure matching and root mean square error between the experimental and observed factor patterns. Several methods of characterizing factor pattern complexity and predicting rotation criterion performance are explored.

14. SUBJECT TERMS
Factor Analysis, Monte Carlo, Varimax, Choleski

15. NUMBER OF PAGES
110

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT
UNCLASSIFIED

20. LIMITATION OF ABSTRACT
UL

END

FILMED

DATE: 4-93

DTIC