

AD-A080 168

GEORGE WASHINGTON UNIV WASHINGTON D C PROGRAM IN LOG--ETC F/G 15/5
USING MULTIVARIATE STATISTICAL ANALYSIS TO OBTAIN READINESS EVA--ETC(U)
NOV 79 Z BARZILY, W H MARLOW, S ZACKS N00014-75-C-0729

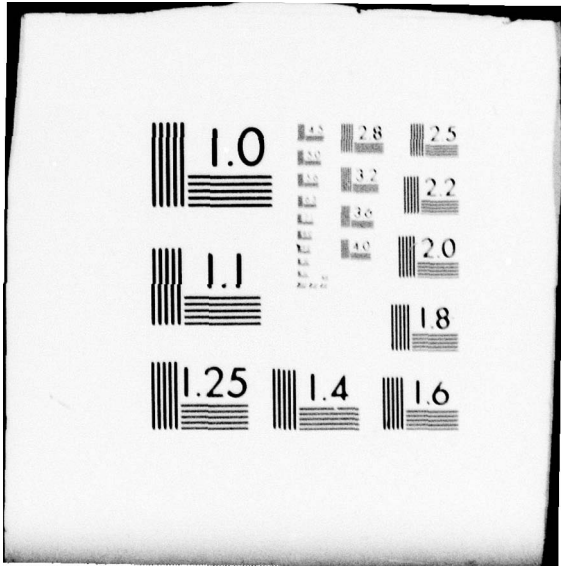
UNCLASSIFIED

SERIAL-T-412

NL

| OF |
AD
A080168

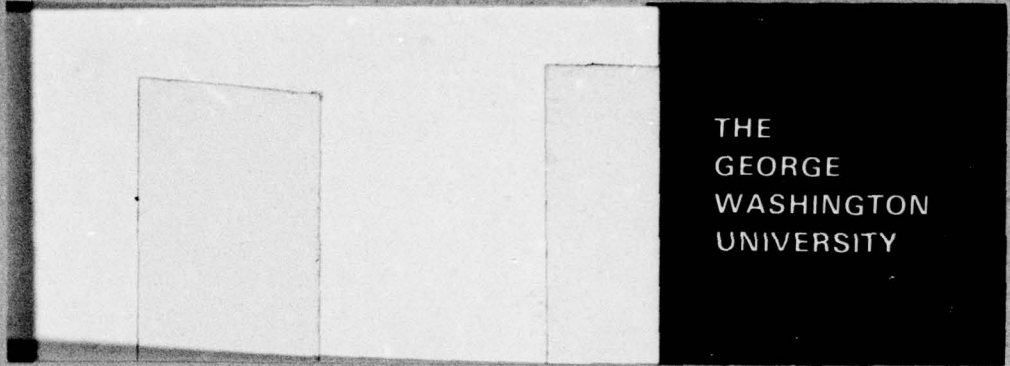




A080168

LEVEL

11/83



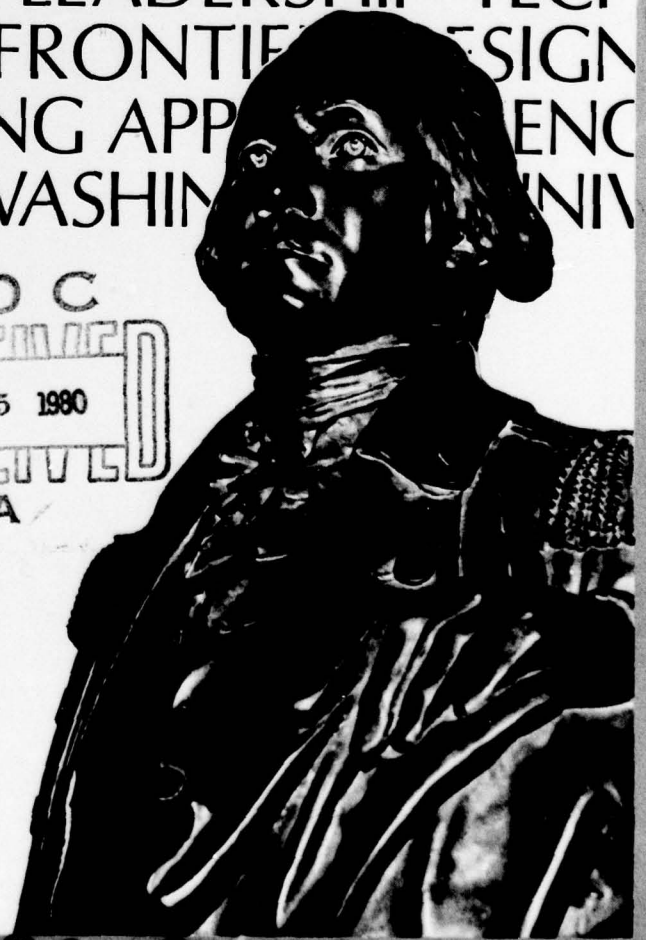
THE
GEORGE
WASHINGTON
UNIVERSITY

STUDENTS FACULTY STUDY R
ESEARCH DEVELOPMENT FUT
URE CAREER CREATIVITY CO
MMUNITY LEADERSHIP TECH
NOLOGY FRONTIER DESIGN
ENGINEERING APP ENO
GEORGE WASHIM UNIV

DDC FILE COPY

DDC
RECEIVED
FEB 5 1980
RECEIVED

A



THIS DOCUMENT IS BEST QUALITY PRACTICABLE.
THE COPY FURNISHED TO DDC CONTAINED A
SIGNIFICANT NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

INSTITUTE FOR MANAGEMENT
SCIENCE AND ENGINEERING
SCHOOL OF ENGINEERING
AND APPLIED SCIENCE

80 2 4 000

DISCLAIMER NOTICE

**THIS DOCUMENT IS BEST QUALITY
PRACTICABLE. THE COPY FURNISHED
TO DDC CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

USING MULTIVARIATE STATISTICAL ANALYSIS
TO OBTAIN READINESS EVALUATIONS

by

Zeev Barzily
W. H. Marlow
S. Zacks

Serial T-412
14 November 1979

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

Program in Logistics
Contract N00014-75-C-0729
Project NR 347 020
Office of Naval Research

This document has been approved for public
sale and release; its distribution is unlimited.

NONE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

<p>14 REPORT DOCUMENTATION PAGE</p>		<p>READ INSTRUCTIONS BEFORE COMPLETING FORM</p>
<p>1. REPORT NUMBER SERIAL T-412</p>	<p>2. GOVT ACCESSION NO. ✓</p>	<p>3. RECIPIENT'S CATALOG NUMBER</p>
<p>4. TITLE (and Subtitle) 6 USING MULTIVARIATE STATISTICAL ANALYSIS TO OBTAIN READINESS EVALUATIONS</p>		<p>5. TYPE OF REPORT & PERIOD COVERED 9 SCIENTIFIC / Rept.)</p> <p>6. PERFORMING ORG. REPORT NUMBER T-412</p>
<p>7. AUTHOR(s) 10 ZEEV BARZILY, W. H. MARLOW S. ZACKS</p>	<p>15 CONTRACT OR GRANT NUMBER(s) N00014-75-C-0729</p>	
<p>9. PERFORMING ORGANIZATION NAME AND ADDRESS THE GEORGE WASHINGTON UNIVERSITY PROGRAM IN LOGISTICS WASHINGTON, D. C. 20037</p>		<p>10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS</p>
<p>11. CONTROLLING OFFICE NAME AND ADDRESS THE OFFICE OF NAVAL RESEARCH CODE 434 ARLINGTON, VIRGINIA 22217</p>		<p>12 REPORT DATE 11 14 November 1979</p>
<p>14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 38</p>		<p>13. NUMBER OF PAGES 35</p>
<p>16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC SALE AND RELEASE; DISTRIBUTION UNLIMITED.</p>		<p>15. SECURITY CLASS. (of this report) NONE</p>
<p>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</p>		<p>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</p>
<p>18. SUPPLEMENTARY NOTES</p>		
<p>19. KEY WORDS (Continue on reverse side if necessary and identify by block number) MULTIVARIATE ANALYSIS READINESS OPERATIONAL READINESS MATERIAL READINESS LOGISTICS</p>		
<p>20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The well-known Statistical Analysis System is used to reduce high-dimensional vectors of data on operational readiness. Such data consist of large numbers of scores for individual ships assigned by experts. The purpose is to provide a robust method of representing the data by a small number of scores that are meaningfully related to the original scores and that allow classification and clustering of the ships on relevant readiness scales.</p>		

DD FORM 1473 1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

405 337

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering
Program in Logistics

Abstract
of
Serial T-412
14 November 1979

USING MULTIVARIATE STATISTICAL ANALYSIS
TO OBTAIN READINESS EVALUATIONS

by

Zeev Barzily
W. H. Marlow
S. Zacks

Accession For	
NTIS GRAB	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	23

The well-known Statistical Analysis System is used to reduce high-dimensional vectors of data on operational readiness. Such data consist of large numbers of scores for individual ships assigned by experts. The purpose is to provide a robust method of representing the data by a small number of scores that are meaningfully related to the original scores and that allow classification and clustering of the ships on relevant readiness scales.

Research supported by
Contract N00014-75-C-0729
Project NR 347 020
Office of Naval Research

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering
Program in Logistics

USING MULTIVARIATE STATISTICAL ANALYSIS
TO OBTAIN READINESS EVALUATIONS

by

Zeev Barzily
W. H. Marlow
S. Zacks

1. Introduction and Summary

Readiness evaluation is one of the most important problem areas in the study of complex military systems. Such studies usually encompass a large number of measurements on the performance of the various subsystems and then attempt to construct reasonable models that relate the evaluation indices of the subsystems in a meaningful functional manner. This is indeed very often a formidable task due to the complexity and multiplicity of variables and functions. However, it is often the case that many of the measured variables correlate with each other. These intercorrelations reveal that variables contain some information on each other. Accordingly, if these intercorrelations can be utilized in a manner that allows considerable reduction in the number of factors to be considered, without much loss in the information in the original data, a significant step can be taken towards simplification of the problem. The present paper applies several well-known multivariate statistical methods to attain this goal. The main objective of the present paper is to discuss what some of the available multivariate statistical methods can attain and to show that such methods can be easily implemented by utilizing appropriate computer packages. In particular, we discuss the methods of principal and

rotated factor analysis and we apply these methods to simulated data on 21 operational readiness variables related to Navy destroyers. The variables and the corresponding parameters were taken from the Institute of Naval Studies study [9]. This study analyzes actual data collected over several years on 83 destroyers. It extends to various aspects of the readiness problem and relates operational readiness to material readiness. As mentioned in our recent survey paper [5], we believe that study [9] is of fundamental importance. It employs a variety of multivariate statistical procedures in a penetrating manner and provides a sound analysis of complicated problems. Our intention is not to duplicate that study but to provide an exposition on the application of the multivariate methods mentioned above. We have chosen to create data sets by simulation and not to use actual data since in this way we can generate data following multivariate normal distributions having specific structures. Thus, by applying the multivariate methods on different sets of simulated data we can illustrate the strength of the methods and what can actually be achieved. We will show that the systems (destroyers) in this example can be classified according to the values of two or three factor scores, which relate all 21 variables in an orthogonal fashion. The factor scores can be graphed and their periodic determination can provide important follow-up on the state of readiness. Statistical control charts can be devised to provide early detection of deterioration in the state of readiness. Similarly, if the data consists of a mixture of two or more samples from different multivariate populations, the plotting of factor scores obtained by a factor analysis of the whole data set can reveal the existence of different clusters. These ideas will be demonstrated in the present paper. We start in Section 2 with a description of the simulations and the structure of the data sets. Section 3 is devoted to principal and rotated factor analysis. In Section 4 we discuss the application of factor analysis to detecting changes in the state of readiness of systems. The mathematical development is presented in appendixes.

The implementation of multivariate analysis of the type discussed in the present paper previously required development of computer programs and systems for data storage and analysis and this hindered its growth and acceptance. Presently there are several statistical computer packages which are available in many computing centers and which are easily implementable. The well-known

Biomedical Computer Programs (P-Series, 1977) [7], Statistical Packages for the Social Sciences (SPSS) [11], and the Statistical Analysis System (SAS) [3] are the most suitable for our purposes. Fortran programs for factor analysis and other related multivariate techniques can be found in Cooley and Lohnes [6], Overall and Klett [10] and more specifically for the methods used in the present paper in our report [4]. In the present paper we apply the SAS procedures. In the appendixes we discuss and also present the SAS programs that we have used.

2. Simulating Data Sets

In the present study we construct data sets on the basis of the operational readiness indices, ORI, of the following 21 variables.

v ₁	Ship control	SHC
v ₂	Navigation	NAV
v ₃	Surface operations - CIC (Combat Information Center)	SOPS
v ₄	Battle communications	BATC
v ₅	Surface gunnery (non-firing)	SGUN
v ₆	AAW (Anti-air Warfare) - CIC	AAWC
v ₇	AAW - Weapons Control	AAWN
v ₈	Engineering	ENG
v ₉	Setting material condition	SMC
v ₁₀	Damage control	DC
v ₁₁	NBC (Nuclear, biological, and chemical)	NBC
v ₁₂	Low-visibility piloting	LVP
v ₁₃	CIC - Assistance in piloting	CICAP
v ₁₄	CIC - Assistance in ASW (Anti-submarine warfare)	CICASW

v_{15}	ECM (Electronic countermeasures)	ECM
v_{16}	Modified full-power run	BFPR
v_{17}	Surface firing	SFIR
v_{18}	AA firing	AAF
v_{19}	Gunfire support	GUNS
v_{20}	Communications	COMM
v_{21}	ASW operations	ASW

The raw scores obtained on these variables by the 83 ships during training can be obtained in the Institute of Naval Studies [9]. We consider rather the ORI's which are indices obtained from the raw scores by the transformation

$$\text{ORI} = 5 + 2 (\text{NSCORE}) \quad (2.1)$$

where NSCORE denotes the standard normal fractile corresponding to the percentile point of the raw score. More precisely, if $x_{(1)} \leq \dots \leq x_{(n)}$ is the order statistic of a sample of n observations on a variable x ,

the NSCORE corresponding to $x_{(i)}$ is $z_{(i)} = \Phi^{-1} \left(\frac{i}{n+1} \right)$, $i = 1, \dots, n$;

where $\Phi(z)$ is the standard normal C.D.F. Theoretically the ORI values, of each of the variables v_i ($i=1, \dots, 21$), are normally distributed with mean $E\{v_i\} = 5$ ($i=1, \dots, 21$) and variance $\text{Var}\{v_i\} = 4$ ($i=1, \dots, 21$). In addition,

the ORI variables, v_i , are not independent. We assume that the vector

$\mathcal{V} = [v_1, \dots, v_{21}]^T$ has a multinormal distribution $N(\mu, \Sigma)$, with mean vector

$\mu = 5\mathbf{1}$, where $\mathbf{1} = [1, \dots, 1]^T$, and covariance matrix $\Sigma = 4\mathcal{R}$ where \mathcal{R}

denotes the matrix of intercorrelations among the 21 variables v_i . For the

purpose of simulating data sets we have used the matrix \mathcal{R} given in the

Institute of Naval Studies [9] and presented here in Table 1. The simulation

TABLE 1
 INTERCORRELATIONS BETWEEN
 ORI VARIABLES v_1 TO v_{21}

1	1.00	0.17	0.27	0.04	0.33	0.31	0.45	0.10	0.06	0.03	0.03	0.32	0.36
	0.06	1.00	0.16	0.07	0.12	0.03	0.01	0.10					
2	0.17	0.06	1.00	0.06	0.18	0.03	0.14	0.22	0.01	0.20	0.04	0.12	0.14
	0.26	0.01	0.07	0.02	0.09	0.09	0.03	0.13					
3	0.27	0.22	0.07	1.00	0.06	0.18	0.03	0.14	0.22	0.01	0.20	0.04	0.14
	0.26	0.01	0.07	0.02	0.09	0.09	0.03	0.13					
4	0.04	0.06	0.06	0.06	1.00	0.23	0.23	0.13	0.11	0.13	0.10	0.03	0.12
	0.03	0.19	0.25	0.03	0.10	0.03	0.04	0.06					
5	0.33	0.18	0.18	0.23	0.23	1.00	0.34	0.05	0.03	0.04	0.01	0.12	0.07
	0.04	0.04	0.25	0.10	0.20	0.03	0.13	0.02					
6	0.31	0.03	0.03	0.23	0.34	0.34	1.00	0.29	0.13	0.13	0.03	0.05	0.03
	0.15	0.06	0.17	0.12	0.13	0.04	0.03	0.10					
7	0.45	0.14	0.14	0.13	0.05	0.29	1.00	0.17	0.17	0.03	0.09	0.14	0.17
	0.11	0.17	0.14	0.01	0.31	0.01	0.12	0.03					
8	0.10	0.22	0.22	0.11	0.03	0.13	0.17	1.00	0.09	0.10	0.00	0.21	0.40
	0.01	0.03	0.10	0.17	0.23	0.12	0.07	0.03					
9	0.05	0.01	0.01	0.13	0.04	0.13	0.17	0.09	1.00	0.13	0.19	0.07	0.14
	0.04	0.14	0.21	0.19	0.05	0.07	0.13	0.13					
10	0.03	0.20	0.20	0.10	0.01	0.03	0.03	0.10	0.13	1.00	0.63	0.06	0.04
	0.07	0.20	0.17	0.03	0.02	0.02	0.29	0.13					
11	0.03	0.04	0.04	0.03	0.12	0.05	0.09	0.00	0.19	0.63	1.00	0.11	0.02
	0.11	0.03	0.09	0.07	0.04	0.13	0.06	0.13					
12	0.32	0.12	0.12	0.12	0.07	0.03	0.14	0.21	0.07	0.06	0.11	1.00	0.00
	0.06	0.18	0.25	0.07	0.07	0.01	0.23	0.21					
13	0.36	0.14	0.14	0.35	0.03	0.22	0.17	0.40	0.14	0.04	0.02	0.00	1.00
	0.12	0.04	0.20	0.11	0.03	0.20	0.17	0.05					
14	0.06	0.26	0.26	0.03	0.04	0.15	0.11	0.01	0.04	0.07	0.11	0.06	0.12
	1.00	0.16	0.29	0.04	0.03	0.06	0.06	0.05					
15	0.17	0.01	0.01	0.19	0.04	0.06	0.17	0.03	0.14	0.20	0.03	0.13	0.04
	0.16	1.00	0.13	0.14	0.14	0.02	0.34	0.51					
16	0.16	0.07	0.07	0.25	0.25	0.17	0.14	0.10	0.21	0.17	0.09	0.25	0.20
	0.29	0.13	1.00	0.01	0.14	0.09	0.23	0.13					
17	0.07	0.02	0.02	0.03	0.10	0.12	0.01	0.17	0.19	0.03	0.07	0.07	0.11
	0.04	0.14	0.01	1.00	0.00	0.05	0.24	0.23					
18	0.12	0.09	0.09	0.10	0.20	0.13	0.31	0.23	0.05	0.02	0.04	0.07	0.03
	0.03	0.14	0.14	0.00	1.00	0.11	0.04	0.12					
19	0.03	0.09	0.09	0.03	0.03	0.04	0.01	0.12	0.07	0.02	0.13	0.01	0.20
	0.06	0.02	0.09	0.05	0.11	1.00	0.12	0.03					
20	0.01	0.03	0.03	0.04	0.18	0.03	0.12	0.07	0.13	0.29	0.06	0.23	0.17
	0.06	0.34	0.23	0.24	0.04	0.12	1.00	0.03					
21	0.10	0.13	0.13	0.06	0.02	0.10	0.03	0.03	0.13	0.13	0.13	0.21	0.05
	0.05	0.51	0.13	0.23	0.12	0.03	0.03	1.00					

was performed according to an algorithm described in Appendix I. It is based on simulating independent standard normal variates, z , and transforming them to corresponding v_i variates ($i=1, \dots, 21$) by a transformation involving the eigenvalues and eigenvectors of the correlation matrix. An SAS program for such simulation is given in Appendix III. In Table 2 we present a sample of $n = 50$ vectors of six variables ($v_1, v_5, v_6, v_7, v_{12}, v_{14}$) simulated according to this program. The sample means, standard deviations (STD DEV) and intercorrelations are provided in Table 3. As illustrated, the sample statistics are generally deviating to some extent (according to their sampling distributions) from the parameters used. However, in actual cases the population parameters are unknown and the analysis must be based exclusively on the sample values, with the possible incorporation of some prior information, and this is what we are doing here.

It should also be remarked that the simulation is based on the matrix of intercorrelations of the above six variables only. This matrix is, however, a submatrix of that given in Table 1 and can be obtained by reading the appropriate rows and columns. In the course of the present study several different data sets were simulated, employing the same algorithm with only slight modifications from case to case, as will be explained later.

3. Principal and Rotated Factor Analysis

It is generally difficult to make comprehensive inference of multivariate data without further analysis, due to the large number of intercorrelated variables. Even in the case of only six variables it will be difficult to discriminate between "good" and "bad" systems, just by inspecting the data sets, or by performing a univariate analysis on each variable separately. The methods of multivariate analysis are designed to provide the needed information in cases of many correlated variables. In the present section we discuss the methods of principal and rotated factor analysis, and show how they can be applied to the evaluation of the readiness of systems. An outline of the theory is given in Appendix II. We refer the reader for an extensive development of the theory and computer programs to the books of Overall and Klett [10], Cooley and Lohnes [5], Tatsuoka [12], and Van de Geer [13].

TABLE 2
 50 SIMULATED VECTORS OF SIX VARIABLES
 ACCORDING TO THE MULTINOMIAL DISTRIBUTION $N(5\mathbf{1}, 4R)$

i	v_1	v_5	v_6	v_7	v_{12}	v_{14}
1	3.503	3.269	7.591	7.124	3.463	5.613
2	4.529	5.294	1.966	5.417	4.021	6.608
3	5.594	5.933	0.347	5.340	6.141	4.522
4	1.891	2.948	7.371	5.303	5.933	5.051
5	2.631	1.105	2.656	1.402	4.078	4.125
6	6.337	4.293	5.396	6.643	6.902	7.504
7	4.963	6.267	4.629	5.431	5.935	2.484
8	2.956	2.448	1.450	1.762	3.117	5.060
9	3.007	3.533	4.034	4.136	4.300	4.747
10	3.951	3.451	3.985	3.774	1.771	3.134
11	5.949	7.307	3.955	6.362	3.373	6.343
12	4.933	4.908	3.409	7.550	4.702	6.852
13	4.244	4.284	2.731	4.367	5.351	4.362
14	3.373	3.360	1.704	4.349	4.340	6.003
15	5.455	2.531	9.759	4.765	4.340	5.939
16	6.076	6.833	6.972	2.700	5.236	7.393
17	2.430	6.347	7.090	1.432	3.639	-1.140
18	4.616	3.535	5.253	7.105	5.440	3.175
19	3.337	4.231	3.646	5.297	5.004	7.421
20	7.915	5.703	6.139	3.344	6.734	10.153
21	4.339	4.453	0.926	2.260	3.113	5.672
22	9.427	6.974	9.020	7.265	7.532	4.959
23	6.122	5.779	5.043	6.474	5.405	3.099
24	7.042	6.516	2.366	7.566	3.579	6.930
25	4.372	6.073	9.149	5.334	2.773	4.261
26	5.922	5.126	4.973	5.062	3.671	3.633
27	2.336	3.512	5.607	0.032	3.345	1.516
28	4.626	3.306	2.232	4.252	7.336	6.023
29	4.007	3.133	5.365	1.313	4.890	1.473
30	6.566	7.713	3.269	7.356	6.933	3.450
31	5.273	3.930	5.021	2.325	2.912	2.715
32	3.079	4.357	3.930	5.727	3.746	7.437
33	4.797	3.643	5.791	5.551	9.067	2.724
34	1.537	-1.307	3.601	3.356	5.196	2.714
35	1.722	2.332	3.334	1.992	1.319	0.916
36	2.967	3.397	1.633	1.195	5.045	2.523
37	4.441	3.769	6.153	1.230	6.165	4.472
38	5.537	5.391	5.139	3.330	0.530	6.516
39	2.534	4.352	5.634	3.576	5.014	1.543
40	5.141	5.169	2.510	2.349	4.054	4.394
41	5.726	2.640	6.141	10.725	3.535	11.300
42	4.643	5.437	5.330	6.513	5.053	7.469
43	7.123	3.347	3.674	3.455	5.530	4.176
44	5.103	7.733	2.357	3.357	4.193	2.373
45	6.176	5.253	5.429	6.230	6.403	3.543
46	0.606	1.079	5.370	3.935	2.371	2.396
47	7.174	4.632	3.537	6.435	4.326	9.936
48	0.549	6.349	5.639	3.537	6.540	2.525
49	3.342	1.917	5.073	4.799	2.103	6.321
50	5.230	6.934	1.036	7.321	6.904	3.135

TABLE 3

SAMPLE STATISTICS OF THE
SIX VARIABLES IN TABLE 2

	v_1	v_5	v_6	v_7	v_{12}	v_{14}
MEANS	4.860	4.577	4.644	4.728	4.993	4.914
STD DEV	1.998	1.897	2.180	2.244	1.916	2.468

CORRELATION MATRIX

	v_1	v_5	v_6	v_7	v_{12}	v_{14}
v_1	1.000000	0.535377	0.183915	0.558162	0.399037	0.519481
v_5	0.535377	1.000000	0.158194	0.362251	0.238443	0.128076
v_6	0.183915	0.158194	1.000000	0.237437	0.192687	-.023431
v_7	0.558162	0.362251	0.237437	1.000000	0.492125	0.617677
v_{12}	0.399037	0.238443	0.192687	0.492125	1.000000	0.239948
v_{14}	0.519481	0.123076	-.023431	0.617677	0.239948	1.000000

3.1 Principal Factor Analysis

The main objective of principal factor analysis is to provide a small number, m , of linear combinations of the original variables v_1, \dots, v_p ($2 \leq m < p$) so that (i) a large proportion of the total variance of the original variables should be accounted for by the m transformed variables, and (ii) the transformed variables should be uncorrelated. It is shown in Appendix II that the solution of this problem is obtained by determining first the m largest eigenvalues of \mathcal{R} and the corresponding eigenvectors; followed by determination of factor scores for each system. Let $\lambda_1 \geq \dots \geq \lambda_p > 0$ be the eigenvalues of the $p \times p$ correlation matrix \mathcal{R} . Since \mathcal{R} is positive definite, these eigenvalues are all real and positive

(with probability one). Moreover $\lambda_1 + \dots + \lambda_p = p$. Hence, choose m so that $(\lambda_1 + \dots + \lambda_m)/p$ is "close enough" to 1. This ratio is the proportion of the sum of variances of $v_i (i=1, \dots, p)$ that is accounted for (explained) by the m factors. These factors are constructed in the following manner. Let $k_{(j)}$ ($j=1, \dots, m$) be the orthonormal eigenvector of R corresponding to λ_j ($j=1, \dots, m$). The m factor-score variables corresponding to $\chi = [v_1, \dots, v_p]^T$ are given by

$$f_j = \frac{1}{\sqrt{\lambda_j}} k_{(j)}^T \chi, \quad j = 1, \dots, m \quad (3.1)$$

where $\chi = [u_1, \dots, u_p]^T$ is a vector of standard scores corresponding to χ , i.e., $u_i = (v_i - \bar{v}_i) / \hat{\sigma}_i$, $i = 1, \dots, p$, \bar{v}_i denotes the sample means of the i th variable and $\hat{\sigma}_i$ designates its sample standard deviation. How much statistical information available in the original vectors of p variables is retained in the m factor scores of the individuals in the sample? To answer this question we introduce additional concepts from the theory of factor analysis.

Consider the matrix ξ , of order $p \times m$, whose m column vectors are related to the m largest eigenvalues and their corresponding eigenvectors, according to the formula:

$$\xi_{(j)} = \lambda_j^{1/2} k_{(j)}, \quad j = 1, \dots, m. \quad (3.2)$$

This matrix is called the factor pattern (structure) matrix. Obviously,

$$\|\xi_{(j)}\|^2 = \lambda_j, \quad j = 1, \dots, m; \text{ and } m = p \text{ then } R = \xi \cdot \xi^T, \text{ or}$$

$$R = \sum_{j=1}^p \xi_{(j)} \xi_{(j)}^T. \text{ This is the spectral decomposition of the correlation}$$

$$\text{matrix } R. \text{ If } m < p \text{ we define } \hat{R}_m = \sum_{j=1}^m \xi_{(j)} \xi_{(j)}^T \text{ and } \tilde{R}_m = R - \hat{R}_m.$$

It is desirable to choose m so that \tilde{R}_m is negligible (or statistically insignificant). Tests of significance of \tilde{R}_m are available (see Cooley and Lohnes [6,103]). According to the spectral decomposition of R , the proportion of the i th diagonal element of R given by the corresponding element of \hat{R}_m is called the communality of the i th variable. It is determined by the formula

$$h_i = \sum_{j=1}^m \lambda_j b_{ij}^2, \quad i = 1, \dots, p. \quad (3.3)$$

One can say that h_i is the proportion of the variance of the i th variable v_i explained by the m factors. In Table 4 we present the eigenvalues and the corresponding eigenvectors of the correlation matrix of Table 3. These eigenvalues and eigenvectors were obtained by employing a computer library routine. On the basis of these values, the first three factor scores were determined for the 50 simulated vectors of Table 2, according to formula (3.1). These factor scores are presented in Table 5.

TABLE 4
EIGENVALUES AND EIGENVECTORS OF THE
CORRELATION MATRIX IN TABLE 3

Eigenvalues					
λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
2.749232	1.068542	0.850006	0.700426	0.384321	0.239473
Eigenvectors					
$k_{(1)}$	$k_{(2)}$	$k_{(3)}$	$k_{(4)}$	$k_{(5)}$	$k_{(6)}$
0.503687	-.035982	0.266720	-.115832	0.707698	0.399522
0.360454	0.292523	0.749462	0.026661	-.329692	-.336753
0.192362	0.755713	-.410358	-.440037	0.072605	-.156839
0.514023	-.118198	-.220703	-.067590	-.608214	0.546450
0.385783	0.141075	-.321752	0.812732	0.117372	-.231162
0.407826	-.555128	-.215730	-.356564	0.038669	-.591894

TABLE 5
FACTOR SCORES OF DATA IN TABLE 2

	f_1	f_2	f_3
1	1.95660	1.44870	0.55230
2	0.03500	-1.25330	0.74550
3	0.24990	-0.94820	1.21970
4	-0.28530	0.73130	-1.92260
5	-1.48420	-0.87110	-0.79760
6	1.09040	-0.13130	-0.38290
7	0.18470	0.30860	0.71440
8	-1.32670	-1.37060	0.10730
9	0.20750	-0.34500	0.23300
10	-0.99640	-0.17310	0.37270
11	0.73150	-0.24330	1.56910
12	0.53350	-0.95340	-0.01960
13	-0.22710	-0.45860	0.09960
14	-0.46930	-1.41490	-0.00790
15	0.15610	1.12330	-1.30390
16	0.56490	0.67710	0.60350
17	-1.15330	2.29530	0.97430
18	0.09070	0.34950	-0.75130
19	0.09630	-0.93620	-0.36110
20	1.97830	-0.59310	-0.63390
21	-0.78450	-1.42940	1.15300
22	1.36510	1.73750	0.05770
23	0.95940	-0.46170	0.05300
24	0.33560	-1.04000	1.26310
25	0.23350	1.65940	0.06770
26	0.00460	0.33000	0.64070
27	-1.55930	1.05760	0.17090
28	0.08350	-0.96790	-0.36100
29	-1.01370	0.93160	-0.23310
30	1.34030	1.95530	0.29340
31	-0.72540	0.44950	0.49450
32	1.33120	-0.56930	-0.30640
33	0.33450	0.97120	-1.26020
34	-1.52690	-0.54330	-2.26140
35	-2.02670	0.02750	0.19030
36	-1.24310	-0.35900	0.62960
37	-0.46100	0.75330	-0.51100
38	-0.34610	-0.32240	1.14530
39	-0.31130	1.13150	-0.13320
40	-0.37360	-0.59930	1.09340
41	1.94010	-1.35010	-2.94330
42	0.64700	-0.08900	-0.33540
43	0.03150	-0.20670	0.31120
44	-0.11050	0.30460	2.19610
45	0.56130	0.66220	0.03220
46	-1.62040	0.23010	-1.55150
47	0.96130	-1.62130	0.07330
48	-0.61170	1.33020	-0.00640
49	-0.66400	-0.85130	-1.10430
50	0.75990	-0.90250	1.17300

Since $(\lambda_1 + \lambda_2 + \lambda_3)/6 = .78$, the three factors explain about 80% of the total variability in the sample. The communalities of the three factors are

<u>Variable</u>	<u>Communality</u>
v_1	.7593
v_5	.9261
v_6	.8551
v_7	.7827
v_{12}	.5184
v_{14}	.8261 .

We see that the three factors explain more than 75% of all variables excluding v_{12} . Notice that the first factor gives more weight to v_1 and v_7 than to the other variables. We can call it therefore the "Ship and AAW Control" factor. Similarly, factor f_2 emphasizes v_6 (AAW-CIC) and gives a large negative weight to v_{14} . This factor can be labeled "Radar and Information Communication." The third factor emphasizes v_5 and deemphasizes v_6, v_7, v_{12} , and v_{14} . It can be labeled "Surface Gunnery." In Figure 1 we present the 50 simulated vectors represented by their first two factor scores. Such a representation can provide a meaningful device for discriminating between "good" and "bad" systems, as graded along the factor scales (f_1, f_2) . Similar scattergrams can be easily provided for (f_1, f_3) and (f_2, f_3) . As will be shown later, such graphical representation of the systems may reveal trends and clusters of subsamples.

Factor analysis of multivariate data can be easily performed by employing available statistical computer packages from SAS, SPSS, or BMDP. We provide in Table 6 the results of an SAS factor analysis procedure performed on 50 simulated vectors from $N(4I, 4R)$, where $p = 6$ and R is

TABLE 6
 FACTOR ANALYSIS WITH SIX VARIABLES,
 PRODUCED BY SAS PROCEDURE ON 50
 SIMULATED VECTORS FROM $N(4I, 4R)$

DEFINITION OF VARIABLES:

COL 1 = v_1 ; COL 2 = v_5 ; COL 3 = v_6 ; COL 4 = v_7 ; COL 5 = v_{12} ; COL 6 = v_{14} .

MEANS AND STD DEVS

	COL 1	COL 2	COL 3	COL 4	COL 5	COL 6
MEAN	3.91766	3.95151	4.49093	4.13207	4.32881	4.06802
STD DEV	1.87219	2.12153	2.02220	2.35387	2.19235	2.18691

CORRELATION MATRIX

	COL 1	COL 2	COL 3	COL 4	COL 5	COL 6
COL 1	1.00000	0.65160	0.13557	0.56443	0.28709	0.57545
COL 2	0.65160	1.00000	0.25144	0.44143	0.21444	0.21628
COL 3	0.13557	0.25144	1.00000	0.12219	0.09678	-0.08323
COL 4	0.56443	0.44143	0.12219	1.00000	0.35855	0.72271
COL 5	0.28709	0.21444	0.09678	0.35855	1.00000	0.23166
COL 6	0.57545	0.21628	0.08323	0.72271	0.23166	1.00000

	1	2	3	4	5	6
EIGENVALUES	2.802324	1.156822	0.846713	0.685932	0.342444	0.165765
PORTION	0.467	0.193	0.141	0.114	0.057	0.028
CUM PORTION	0.467	0.660	0.801	0.915	0.972	1.000

EIGENVECTORS

	1	2
COL 1	0.51049	0.04610
COL 2	0.41630	0.38367
COL 3	0.12621	0.78333
COL 4	0.50975	-0.17007
COL 5	0.29683	0.03120
COL 6	0.44968	-0.45516

FACTOR PATTERN

	FACTOR 1	FACTOR 2
COL 1	0.85457	0.04958
COL 2	0.69688	0.41266
COL 3	0.21127	0.84251
COL 4	0.85334	-0.18292
COL 5	0.49690	0.03355
COL 6	0.75277	-0.48956

FINAL COMMUNALITY ESTIMATES:

COL 1	COL 2	COL 3	COL 4	COL 5	COL 6
0.732752	0.655935	0.754461	0.761642	0.248036	0.806322

the sample correlation matrix of Table 3. The results obtained are similar to those presented earlier. For the purpose of simulating the multivariate data sets and applying factor analysis on the simulated data, we have found that the SAS package is most convenient. With the features available on SAS we could conveniently execute simulation and factor analysis and other statistical procedures in one program (see Appendix III).

3.2 Rotated Factor Analysis

One of the main problems of principal factor analysis is that the m factor score variables (3.1) are often linear combinations which ascribe high relative weight to many of the original variables, and no immediate (or direct) interpretation can be given to the factor scores. In order to obtain factor scores which depend on a small number of the original variables various rotational methods have been developed (see Harman [8]), which yield vectors of factors coefficients with a large number of elements close to zero. We consider here orthogonal transformations which reduce the pattern matrix \mathcal{S} to a rotated factor space matrix $\mathcal{A} = \mathcal{S} \cdot \mathcal{R}$, where \mathcal{R} is an $m \times m$ orthogonal matrix determined so that the column vectors of \mathcal{A} have as many zero entries as possible. Statistics packages provide various options for orthogonal and oblique rotation of the factor space. In Table 7 we show the result of EQUAMAX of orthogonal rotation of the factor pattern matrix of Table 6. This rotation is designed to maximize an adjusted fourth moment of the elements of the column vectors of the resulting factor pattern matrix \mathcal{A} (see Harman [8, p. 299]). Another commonly applied rotation is called VARIMAX, which maximizes the fourth moment (unadjusted) of each column of \mathcal{A} . Both methods of rotation are frequently applied without yielding significant difference in the results. For a comparison of various orthogonal rotations, including the VARIMAX and EQUAMAX see Harman [8]. Notice that the factor scores corresponding to the rotated factor analysis are obtained by multiplying the standardized individual vectors \mathcal{Z} by the matrix, $\mathcal{S} \mathcal{A}^{-1} \mathcal{R}$. The "Rotated Factor Pattern" matrix in Table 7 is the matrix $\mathcal{A} = \mathcal{S} \mathcal{R}$, where \mathcal{S} is the "Factor Pattern" matrix of Table 6 and \mathcal{R} is the "Orthogonal Transformation Matrix"

TABLE 7

EQUAMAX ROTATION OF FACTOR PATTERN MATRIX
RELATED TO EXAMPLE OF TABLE 6

ROTATION METHOD: EQUAMAX

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2
COL1	0.30537	0.29005
COL2	0.55114	0.59345
COL3	-0.03650	0.86783
COL4	0.37016	0.06675
COL5	0.46695	0.17313
COL6	0.36074	-0.25532

ORTHOGONAL TRANSFORMATION MATRIX

	1	2
1	0.95339	0.28377
2	-0.28377	0.95339

PROPORTIONAL CONTRIBUTIONS TO COMMON VARIANCES BY ROTATED FACTORS

FACTOR1	FACTOR2
2.669317	1.289329

SCORING COEFFICIENT MATRIX

	FACTOR1	FACTOR2
COL1	0.28025	0.12763
COL2	0.13723	0.41262
COL3	-0.13433	0.71975
COL4	0.33636	-0.06521
COL5	0.16130	0.07813
COL6	0.37767	-0.32957

of Table 7. The "Scoring Coefficient Matrix" of Table 7 is equal to $SA^{-1}R$. Its two columns provide the coefficients with which to multiply the individual vectors Z to obtain their factor scores. Inspection of the scoring coefficients in Table 7 shows that the first (rotated) factor emphasizes variables v_1 , v_7 , and v_{14} while the second (rotated) factor emphasizes variables v_5 and v_6 . Variable v_{12} does not attain a considerable weight in this representation. We further illustrate the method of rotated factor analysis by performing such an analysis on all 21 ORI variables. In Tables 8-10 we present the results of such an analysis with EQUAMAX rotation. The data set consists of 50 simulated vectors of 21 components following the distribution $N(5I, 4R)$, where R is the correlation matrix of Table 1.

In the present analysis we display only the first three principal factors and their orthogonal rotation. As seen in Table 8, the first three principal factors account for only 38.7% of the total variability. In order to account for 80% of the variability we have to retain ten principal factors. This is not surprising, in light of the rather small intercorrelations between many of the 21 ORI variables (see Table 1). It is very difficult to ascribe meaning to the principal factors without rotation (see Table 9). However, after an EQUAMAX rotation we obtain factor-scores coefficients which can provide relevant interpretation to the factors (Table 10), although this interpretation is different from the one obtained by analyzing six variables only. Thus, (rotated) factor 1 emphasizes variables v_{13} , v_{14} , v_{16} and to some extent also v_2 , v_3 , v_{19} , v_{20} . Most of these variables relate to different aspects of navigation, piloting, and anti-submarine warfare. (Rotated) factor 2 emphasizes v_4 , v_6 , v_{10} , v_{11} and deemphasizes v_{20} , (anti-air warfare), engineering, and damage control. (Rotated) factor 3 emphasizes v_9 , v_{15} , and v_{21} , which relate to electronic operations and setting material conditions. As explained earlier, not all the aspects of the operational readiness are represented by the three rotated factors. One needs about ten rotated factors to account for a large portion of the variability in 21 variables.

TABLE 8

EIGENVALUES AND EIGENVECTORS OF
CORRELATION MATRIX IN TABLE 1

	1	2	3	4	5	6
EIGENVALUES	3.794984	2.415706	1.915619	1.732022	1.520571	1.319101
PORTION	0.181	0.115	0.091	0.082	0.072	0.063
CUM PORTION	0.181	0.296	0.387	0.469	0.542	0.605
	7	8	9	10	11	12
EIGENVALUES	1.285502	1.212352	0.955083	0.833270	0.756898	0.633716
PORTION	0.061	0.058	0.045	0.040	0.036	0.030
CUM PORTION	0.666	0.724	0.769	0.809	0.845	0.875
	13	14	15	16	17	18
EIGENVALUES	0.597649	0.460369	0.414631	0.361732	0.289878	0.193707
PORTION	0.023	0.022	0.020	0.017	0.014	0.009
CUM PORTION	0.903	0.925	0.945	0.962	0.976	0.985
	19	20	21			
EIGENVALUES	0.147695	0.115658	0.043856			
PORTION	0.007	0.006	0.002			
CUM PORTION	0.992	0.998	1.000			

EIGENVECTORS

	1	2	3
COL 1	0.19064	-0.02192	0.00061
COL 2	0.31186	-0.00371	-0.07639
COL 3	0.34164	-0.01266	0.00310
COL 4	0.22031	-0.23721	0.18299
COL 5	0.25561	0.15462	-0.00189
COL 6	0.14146	-0.26245	0.23932
COL 7	0.13558	-0.28916	0.06381
COL 8	0.31212	0.04438	0.17556
COL 9	0.20694	0.22207	0.34169
COL 10	0.21507	-0.07917	0.27990
COL 11	0.17075	-0.25016	0.26444
COL 12	0.25605	0.24543	-0.01065
COL 13	0.24580	-0.17586	-0.26415
COL 14	0.22924	-0.13077	-0.32797
COL 15	0.02812	0.36811	0.17200
COL 16	0.31317	0.04039	-0.34390
COL 17	0.10458	0.29794	-0.03943
COL 18	0.23165	0.08198	-0.03293
COL 19	0.16118	-0.16203	-0.16226
COL 20	0.07772	0.33043	-0.41069
COL 21	0.08802	0.41257	0.28378

TABLE 9
COMMUNALITIES, FACTOR PATTERN MATRIX, AND TRANSFORMATION
MATRIX FOR FACTOR ANALYSIS OF TABLE 1

FINAL COMMUNALITY ESTIMATES:						
COL1	COL2	COL3	COL4	COL5	COL6	COL7
0.139092	0.380294	0.443340	0.384266	0.305708	0.352044	0.279543
COL8	COL9	COL10	COL11	COL12	COL13	COL14
0.433502	0.505296	0.340764	0.395783	0.394545	0.437672	0.446793
COL15	COL16	COL17	COL18	COL19	COL20	COL21
0.337011	0.602679	0.258913	0.221960	0.212452	0.609789	0.594864

FACTOR PATTERN			
	FACTOR1	FACTOR2	FACTOR3
COL1	0.37139	-0.03407	0.00085
COL2	0.60752	-0.00577	-0.10572
COL3	0.66553	-0.01963	0.00429
COL4	0.42917	-0.36869	0.25327
COL5	0.49795	0.24032	-0.00262
COL6	0.27557	-0.40791	0.33123
COL7	0.26412	-0.44943	0.08832
COL8	0.60803	0.06897	0.24299
COL9	0.40313	0.34516	0.47292
COL10	0.41898	-0.12305	0.38740
COL11	0.33264	-0.38882	0.36600
COL12	0.49831	0.38146	-0.01474
COL13	0.47884	-0.27334	-0.36561
COL14	0.44658	-0.20325	-0.45393
COL15	0.05473	0.57214	0.23805
COL16	0.61007	0.06278	-0.47597
COL17	0.20372	0.46307	-0.05458
COL18	0.45127	0.12742	-0.04558
COL19	0.31400	-0.25183	-0.22458
COL20	0.15140	0.51357	-0.56843
COL21	0.17147	0.64124	0.39277

ORTHOGONAL TRANSFORMATION MATRIX

	1	2	3
1	0.75536	0.43159	0.49235
2	-0.14824	-0.61965	0.77076
3	-0.63773	0.65557	0.40438

PROPORTIONAL CONTRIBUTIONS TO COMMON VARIANCES BY ROTATED FACTORS

FACTOR1	FACTOR2	FACTOR3
3.000329	2.457702	2.668277

TABLE 10
 ROTATED FACTOR MATRIX AND FACTOR-SCORES COEFFICIENTS
 FOR FACTOR ANALYSIS OF TABLE 1

ROTATED FACTOR PATTERN			
	FACTOR1	FACTOR2	FACTOR3
COL1	0.28523	0.18196	0.15693
COL2	0.52743	0.19647	0.25191
COL3	0.50323	0.30224	0.31425
COL4	0.21753	0.57972	0.02955
COL5	0.34242	0.06428	0.42933
COL6	0.05752	0.53884	-0.04473
COL7	0.20994	0.45037	-0.13064
COL8	0.29440	0.37398	0.45079
COL9	-0.04306	0.27014	0.65575
COL10	0.03737	0.51104	0.26811
COL11	0.07566	0.62443	0.01210
COL12	0.32938	-0.03075	0.53365
COL13	0.63562	0.13636	-0.12276
COL14	0.65716	0.02110	-0.12035
COL15	-0.19523	-0.17482	0.56421
COL16	0.75536	-0.03763	0.15628
COL17	0.12015	-0.23479	0.43514
COL18	0.35128	0.08593	0.30196
COL19	0.41789	0.14434	-0.13032
COL20	0.40031	-0.62553	0.24052
COL21	-0.21593	-0.06535	0.73750

SCORING COEFFICIENT MATRIX			
	FACTOR1	FACTOR2	FACTOR3
	0.07573	0.05127	0.03749
	0.15655	0.03439	0.05466
	0.13233	0.03220	0.03097
	0.02379	0.23005	-0.00349
	0.03530	-0.00591	0.14073
	-0.03035	0.24933	-0.02447
	0.05073	0.17554	-0.09043
	0.03593	0.13461	0.15219
	-0.09333	0.11915	0.26226
	-0.03797	0.21179	0.09633
	-0.03173	0.26232	-0.00364
	0.03035	-0.04616	0.18331
	0.23336	-0.00055	-0.10227
	0.25254	-0.05242	-0.10274
	-0.10345	-0.05906	0.23991
	0.27611	-0.10961	-0.00130
	0.03033	-0.11429	0.16266
	0.09724	0.00304	0.03953
	0.15276	0.02345	-0.03702
	0.13733	-0.30904	0.06351
	-0.13596	-0.01057	0.30975

4. Detecting Deterioration in Readiness and Subgroups

We have seen in the previous section that the readiness of systems can be represented by principal factors or rotated factors. This is a combined measurement of readiness, which transforms the basic ORI scores and reduces them to a small number of orthogonal factor scores. This representation of the readiness of systems is particularly useful for control purposes. Suppose that we wish to follow the state of readiness of a particular system. We can periodically make observations on the ORI variables and present the corresponding factor scores on the scattergrams similar to the one in Figure 1. Significant deterioration in readiness will be detected by the location of these points in the scattergram. Moreover, if a whole group of points cluster on the scattergram on the negative side of a factor there may be an indication that this group originates from a different population and further analysis should follow. Such a case is illustrated in Figure 2, in which the factor scores obtained by a rotated factor analysis of the 6 ORI variables, when the sample of 50 systems consisted of 25 units from the distribution $N(5\bar{\mu}, 4\bar{\sigma})$ and 25 units from the distribution $N(\bar{\mu}, 4\bar{\sigma})$. The points in Figure 2 corresponding to the units in the first subsample are labeled "1" and the others are labeled "2". It is seen that most of the second subsample points are concentrated at the negative part of f_1 .

There is a strong indication of a significant difference between the two subsamples. The capability of rotated factor analysis to separate such subsamples in the new factor space is not surprising. It can be given precise algebraic and geometric interpretations. We do not dwell on this here to any further degree but only remark that if such separation of two natural subsamples (as two different subfleets) is indicated then one should reinforce the analysis by performing another method of multivariate analysis, which is designed for discrimination between subgroups and classification of the individual units to various readiness groups according to their distances from the centroids of these groups. For the theory and explanation of these methods see Tatsuoka [12], Afifi and Azen [1], and Van De Geer [13].

For the application of a stepwise Discriminant Analysis procedure, employing an SPSS program on the simulated data with 21 ORI variables, see our report [4].

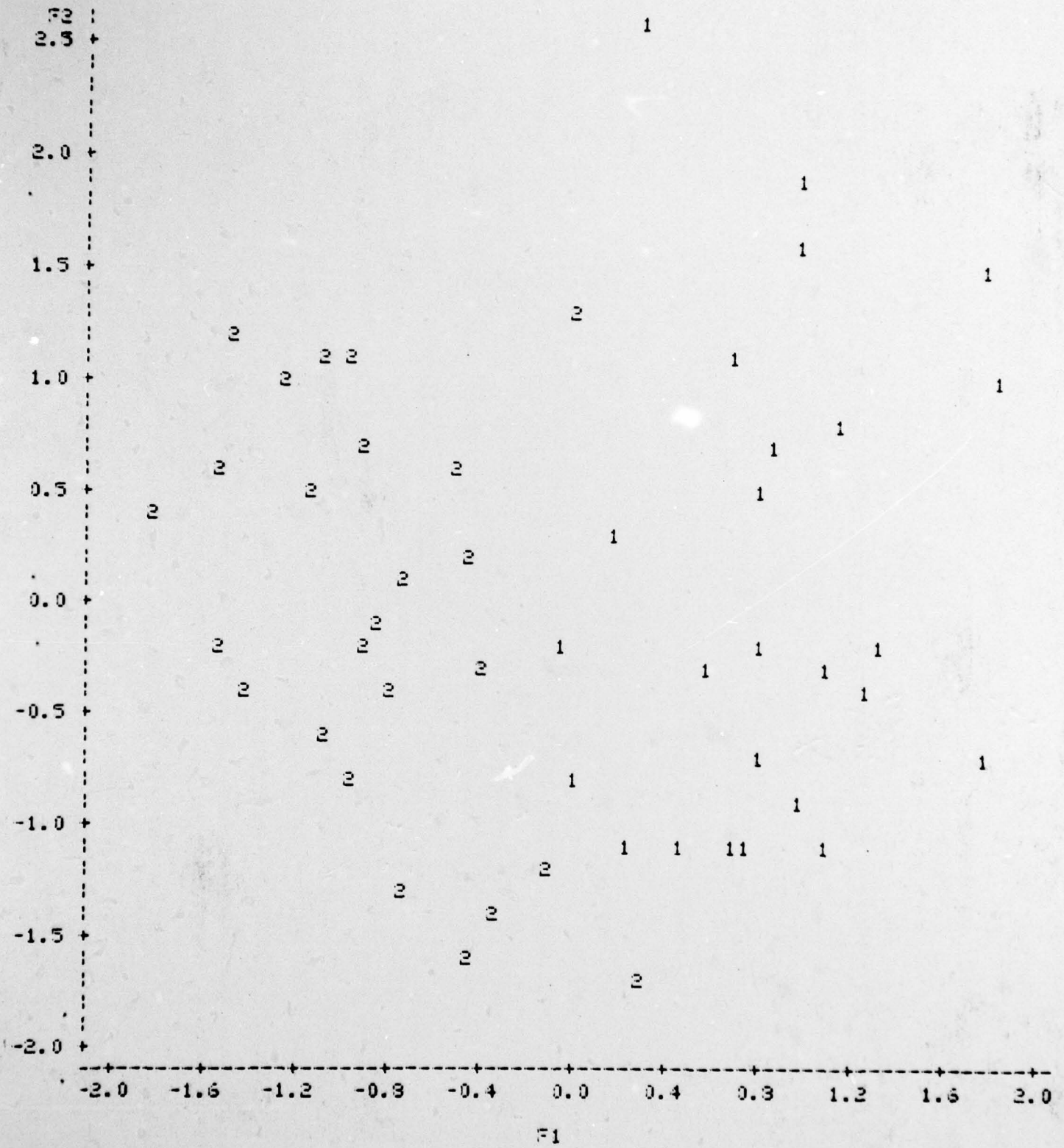


Figure 2. Scattergram of f_2 versus f_1 in mixed samples.

REFERENCES

- [1] AFIFI, A. A. and S. P. Azen (1972). Statistical Analysis: A Computer Oriented Approach. Academic.
- [2] ANDERSON, T. W. (1958). Introduction to Multivariate Analysis. Wiley.
- [3] BARR, A. J., S. H. GOODNIGHT, J. P. SALL, and J. T. HELWIG (1979). A User's Guide to SAS79. SAS Institute, Raleigh, N. C.
- [4] BARZILY, Z., W. H. MARLOW, and S. ZACKS (1978). Readiness evaluation using multivariate data reduction. Technical Paper Serial T-385. Program in Logistics, The George Washington University.
- [5] BARZILY, Z., W. H. MARLOW, and S. ZACKS (1979). Survey of approaches to readiness. Naval Res. Logist. Quart. 26 21-31.
- [6] COOLEY, W. W. and P. R. LOHNES (1971). Multivariate Data Analysis. Wiley.
- [7] DIXON, W. J. (Ed.) (1977). BMDP - Biomedical Computer Programs. University of California Press.
- [8] HARMAN, H. H. (1976). Modern Factor Analysis. 3rd ed. (revised). University of Chicago Press.
- [9] LOCKMAN, R. F., P. H. STOLOFF, B. H. MANHEIMER, J. R. HARDGRAVE, and W. F. STORY (1970). Navy Readiness Analysis System Methodology Study. 1 Institute of Naval Studies of the Center for Naval Analyses. Study No. 27. University of Rochester.

- [10] OVERALL, J. and C. T. KLETT (1972). Applied Multivariate Analysis. McGraw-Hill.
- [11] NORMAN, H. N., C. H. HULL, J. G. JENKINS, K. STEINBRENNER, and D. H. BENT (1975). SPSS: Statistical Package for the Social Sciences. 2nd ed. McGraw-Hill.
- [12] TATSUOKA, M. M. (1971). Multivariate Analysis: Techniques for Educational and Psychological Research. Wiley.
- [13] VAN DE GEER, J. P. (1971). Introduction to Multivariate Analysis for the Social Sciences. Freeman.

APPENDIX I

SIMULATING MULTINORMAL VECTORS

In the present appendix we discuss use of the SAS computer package for simulation of p -dimensional multinormal vectors. PROCEDURE MATRIX of SAS provides the possibility of computing the eigenvalues and an orthonormal matrix of eigenvectors of the correlation matrix R . This means that the simulation can be based on the following result from the theory of multinormal distributions.

Let R be an orthonormal matrix of the eigenvectors of a correlation matrix, R , and let Λ be a diagonal matrix whose elements are the eigenvalues of R . Assume that R is of full rank. We can write $R = R\Lambda R^T$. Accordingly, let $C = R\Lambda^{1/2}$ and let Z be a p -dimensional vector of independent standard normal random variables, that is, $Z \sim N(0, I)$, then the distribution of $\mu = CZ$ is $N(0, R)$. Finally, $X = \mu + 2\mu$ is distributed like $N(\mu, 4R)$. The SAS program given in Table 1, Appendix III simulates 50 independent 21-dimensional multinormal vectors with mean vector 5μ and covariance matrix $4R$, where R is the correlation matrix of Table 1 and it then performs the rotated factor analysis. The program is based on the following algorithm:

- Step 1. Read R ;
- Step 2. Generate $N = 50$ vectors of 21 independent standard normal random variables;
- Step 3. Determine the eigenvalues and normalized eigenvectors of R ;
- Step 4. Arrange the eigenvalues in a diagonal matrix Λ and the eigenvectors in a matrix R ;
- Step 5. Determine $S = \Lambda^{1/2}$ and $C = R \cdot S$;
- Step 6. Arrange the data generated in Step 2 in a 50×21 matrix X ;

Step 7. Make the transformation $W = 2 \cdot Y \cdot C^T$:

Step 8. Determine the matrix $M = 5 \cdot J$, where J is a 50×21 matrix of 1's;

Step 9. Compute $X^* = X + M$.

The matrix X^* consists of 50 i.i.d. row vectors, each of which is distributed like $N(5\mu, 4\sigma)$.

If a SAS package is not available one can perform the simulation by another method which does not require the determination of eigenvalues and eigenvectors but only the solution of linear equations. A FORTRAN program of such a procedure, based on a recursive algorithm, is given in our report [4].

APPENDIX II

PRINCIPAL AND ROTATED FACTOR ANALYSIS

Let $\mu \sim N(0, R)$ be a standard multivariate normal vector. The distribution of $\ell^T \mu$ is like that of $N(0, \ell^T R \ell)$. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ be the eigenvalues of R . We wish to determine a vector (functional) ℓ , with length $|\ell| = 1$, which maximizes the variance of $\ell^T \mu$. The Lagrangian is

$$f(\ell, \lambda) = \ell^T R \ell - \lambda(\ell^T \ell - 1) \quad (\text{A.2.1})$$

Differentiating $f(\ell, \lambda)$ with respect to ℓ yields the eigenstructure equation

$$R \ell = \lambda \ell \quad (\text{A.2.2})$$

Notice that

$$\ell^T R \ell = \lambda \ell^T \ell = \lambda$$

Thus, $\ell^{(1)}$ is an eigenvector of R , of unit length, corresponding to the largest eigenvalue of R , namely to λ_1 which is the variance of $\ell^{(1)T} \mu$.

Similarly, let $\ell^{(2)}, \dots, \ell^{(p)}$ be the eigenvectors of unit length of R , corresponding to $\lambda_2, \dots, \lambda_p$. Notice that the variance of $\ell^{(i)T} \mu$ is λ_i ($i=1, \dots, p$) and that

$$\text{cov}(\ell^{(i)T} \mu, \ell^{(j)T} \mu) = 0, \quad \text{all } i \neq j. \quad (\text{A.2.2})$$

Indeed, if $R_i = \lambda_i \ell^{(i)} \ell^{(i)T}$, $i = 1, \dots, p$, then the spectral decomposition of R is

$$R = \sum_{j=1}^p \lambda_j \ell^{(j)} \ell^{(j)T} \quad (\text{A.2.3})$$

Furthermore, for any $i \neq j$

$$\begin{aligned} \text{cov}(\xi^{(i)T} \mu, \xi^{(j)T} \mu) &= \xi^{(i)T} R \xi^{(j)} \\ &= \sum_{k=1}^p \xi^{(i)T} R_{kk} \xi^{(j)} = 0 \end{aligned} \quad (\text{A.2.4})$$

Let $R = (\xi^{(1)}, \dots, \xi^{(p)})$ be an orthogonal matrix with columns which are the eigenvectors of R . The distribution of

$$\xi = \Lambda^{-1/2} R^T \mu \quad (\text{A.2.5})$$

is like that of $N(Q, \Lambda)$; where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$. Indeed,

$R^T \mu \sim N(Q, R^T R R R^T)$. But $R^T R R R^T = \Lambda$. The components of ξ are called the principal factor scores, corresponding to μ . The orthogonal transformation of μ , given by (A.2.5) yields independent standard normal random variables.

Since $\text{trace } R = \text{trace } \Lambda = \sum_{i=1}^p \lambda_i = p$, the ratio λ_i/p ($i=1, \dots, p$)

is the proportion of the total variance of μ accounted for by

f_i ($i=1, \dots, p$). If we choose only the first m ($1 \leq m < p$) eigenvectors of

R , corresponding to $\lambda_1 \geq \dots \geq \lambda_m$, and define $R_{(m)} = [\xi^{(1)}, \dots, \xi^{(m)}]$,

$\Lambda_{(m)} = \text{diag}(\lambda_1, \dots, \lambda_m)$, the transformation $\Lambda_{(m)}^{-1/2} R_{(m)}^T \mu$ yields the first

m components of ξ . The concepts of communality and the nature of rotated factor analysis was explained in Section 3.

APPENDIX III
COMPUTER PROGRAMS

In the present appendix we present two SAS programs. The program in Table III.1 performs simulation of 50 independent multinormal vectors of 21 components, having the common distribution $N(5\lambda, 4R)$. A rotated factor analysis is then performed on the simulated data set with an EQUAMAX rotation. In Table III.2 we present a program according to which a scattergram of the factor scores, corresponding to the simulated data set, can be obtained. This program is designed to present the scattergram of f_2 versus f_1 in the case of two subsamples of size $n = 25$ from $N(4\lambda, 4R)$ and $N(2.5\lambda, 4R)$, respectively. Figure 2 was obtained by a similar program, with two subsamples from $N(5\lambda, 4R)$ and $N(\lambda, 4R)$. The points from the two subsamples are labeled "1" and "2", respectively. The first part of the program simulates 50 6-dimensional normal vectors from these two distributions. The factor weights, to obtain f_1 and f_2 are read into the program as a data set ("DATA FSCORE"). These weights can be obtained by performing the factor analysis program in Table III.1.

TABLE III.1

SAS PROGRAM FOR SIMULATING 50 $N(5\lambda, 4R)$ VECTORS AND
 PERFORMING ROTATED FACTOR ANALYSIS ON 21 VARIABLES

```

OPTIONS LS=30;
DATA CORR;
INPUT C1-C21;
TITLE FACTOR ANALYSIS FOR READINESS;
CARDS;
1.00 0.17 0.27 0.04 0.38 0.31 0.45 0.10 0.06 0.03 0.03 0.32 0.36
0.06 0.17 0.16 0.07 0.12 0.03 0.01 0.10
0.17 1.00 0.22 0.06 0.18 0.03 0.14 0.22 0.01 0.20 0.04 0.12 0.14
0.26 0.01 0.07 0.02 0.09 0.09 0.03 0.13
0.27 0.22 1.00 0.06 0.18 0.03 0.14 0.22 0.01 0.20 0.04 0.12 0.14
0.26 0.01 0.07 0.02 0.09 0.09 0.03 0.13
0.04 0.06 0.06 1.00 0.23 0.23 0.13 0.11 0.13 0.10 0.03 0.12 0.35
0.03 0.19 0.25 0.03 0.10 0.03 0.04 0.06
0.38 0.18 0.18 0.23 1.00 0.34 0.05 0.03 0.04 0.01 0.12 0.07 0.03
0.04 0.04 0.25 0.10 0.20 0.03 0.18 0.02
0.31 0.03 0.03 0.23 0.34 1.00 0.29 0.13 0.13 0.03 0.05 0.03 0.22
0.15 0.06 0.17 0.12 0.13 0.04 0.03 0.10
0.45 0.14 0.14 0.13 0.05 0.29 1.00 0.17 0.17 0.03 0.09 0.14 0.17
0.11 0.17 0.14 0.01 0.31 0.01 0.12 0.03
0.10 0.22 0.22 0.11 0.03 0.13 0.17 1.00 0.09 0.10 0.00 0.21 0.40
0.01 0.03 0.10 0.17 0.28 0.12 0.07 0.03
0.06 0.01 0.01 0.13 0.04 0.13 0.17 0.09 1.00 0.18 0.19 0.07 0.14
0.04 0.14 0.21 0.19 0.05 0.07 0.13 0.13
0.03 0.20 0.20 0.10 0.01 0.03 0.03 0.10 0.18 1.00 0.68 0.06 0.04
0.07 0.20 0.17 0.03 0.02 0.02 0.29 0.18
0.03 0.04 0.04 0.03 0.12 0.05 0.09 0.00 0.19 0.68 1.00 0.11 0.02
0.11 0.03 0.09 0.07 0.04 0.13 0.06 0.13
0.32 0.12 0.12 0.12 0.07 0.03 0.14 0.21 0.07 0.06 0.11 1.00 0.00
0.06 0.18 0.25 0.07 0.07 0.01 0.23 0.21
0.36 0.14 0.14 0.35 0.03 0.22 0.17 0.40 0.14 0.04 0.02 0.00 1.00
0.12 0.04 0.20 0.11 0.03 0.20 0.17 0.05
0.06 0.26 0.26 0.03 0.04 0.15 0.11 0.01 0.04 0.07 0.11 0.06 0.12
1.00 0.16 0.29 0.04 0.03 0.06 0.06 0.05
0.17 0.01 0.01 0.19 0.04 0.06 0.17 0.03 0.14 0.20 0.03 0.13 0.04
0.16 1.00 0.13 0.14 0.14 0.02 0.34 0.51
0.16 0.07 0.07 0.25 0.25 0.17 0.14 0.10 0.21 0.17 0.09 0.25 0.20
0.29 0.13 1.00 0.01 0.14 0.09 0.28 0.13
0.07 0.02 0.02 0.03 0.10 0.12 0.01 0.17 0.19 0.03 0.07 0.07 0.11
0.04 0.14 0.01 1.00 0.00 0.05 0.24 0.23
0.12 0.09 0.09 0.10 0.20 0.13 0.31 0.28 0.05 0.02 0.04 0.07 0.03
0.03 0.14 0.14 0.00 1.00 0.11 0.04 0.12
0.03 0.09 0.09 0.03 0.03 0.04 0.01 0.12 0.07 0.02 0.13 0.01 0.20
0.06 0.02 0.09 0.05 0.11 1.00 0.12 0.03
0.01 0.03 0.03 0.04 0.13 0.03 0.12 0.07 0.13 0.29 0.06 0.23 0.17
0.06 0.34 0.28 0.24 0.04 0.12 1.00 0.03
0.10 0.13 0.13 0.06 0.02 0.10 0.03 0.03 0.18 0.18 0.13 0.21 0.05
0.05 0.51 0.13 0.23 0.12 0.03 0.03 1.00

```

TABLE III.1 (Cont'd)

```

DATA RAND;
KEEP X1-X21;
LOOP:N+1;
X1=NORMAL(0);X2=NORMAL(0);X3=NORMAL(0);X4=NORMAL(0);X5=NORMAL(0);X6=NORMAL(0);
X7=NORMAL(0);X8=NORMAL(0);X9=NORMAL(0);X10=NORMAL(0);
X11=NORMAL(0);X12=NORMAL(0);X13=NORMAL(0);X14=NORMAL(0);X15=NORMAL(0);
X16=NORMAL(0);X17=NORMAL(0);X18=NORMAL(0);X19=NORMAL(0);X20=NORMAL(0);
X21=NORMAL(0);
OUTPUT;
IF N < 50 THEN GO TO LOOP;
PROC MATRIX;
FETCH R DATA=CORR;
EIGEN M E R;
D=DIAG(M);
S=SQRT(D);
B=E*S;
FETCH Y DATA=RAND;
W=Y*B;
W=2*W;
IA=1:50;W1=W(IA,);MW1=J(50,21,5);
W=W1+MW1;
OUTPUT W OUT=DISC(KEEP=COL1-COL21);
PROC FACTOR NFACT=3 OUT=FACT1 METHOD=PRIN ROTATE=EQUAMAX EIGENVECTORS SCORE;
VAR COL1-COL21;

```

TABLE III.2

SIMULATING MULTINORMAL VECTORS WITH SIX COMPONENTS FROM TWO
DISTRIBUTIONS AND PLOTTING THEIR FACTOR SCORES

```

OPTIONS LS=30;
DATA CORR;
INPUT C1-C6;
TITLE FACTOR ANALYSIS FOR READINESS;
CARDS;
1.0000 0.5354 0.1839 0.5582 0.3390 0.5195
0.5354 1.0000 0.1582 0.3623 0.2384 0.1281
0.1839 0.1582 1.0000 0.2374 0.1927 -.0234
0.5582 0.3623 0.2374 1.0000 0.4921 0.6177
0.3390 0.2384 0.1927 0.4921 1.0000 0.2399
0.5195 0.1281 -.0234 0.6177 0.2399 1.0000
DATA RAND;
KEEP X1-X6;
  LOOP=N+1;

X1=NORMAL(0);X2=NORMAL(0);X3=NORMAL(0);X4=NORMAL(0);X5=NORMAL(0);X6=NORMAL(0);

OUTPUT;
IF N < 50 THEN GO TO LOOP;
DATA FSCORE;INPUT F1-F2;CARDS;
0.49979 -.014726
0.52570 -.024989
-.21200 0.48125
0.14143 .27133
-.13526 0.50603
0.17949 0.17796
PROC MATRIX;
FETCH R DATA=CORR;
EIGEN M E R;
D=DIAG(M);
S=SQRT(D);
B=E*S;
FETCH Y DATA=RAND;
W=Y*B';
W=2*W;
IA=1:25;W1=W(IA,♦);MW1=J(25,6,4);
W1=W1+MW1;
IB=26:50;W2=W(IB,♦);MW2=J(25,6,2.5);W2=W2+MW2;
W=W1//W2;
ID=I(50);W3=J(50,50,1);W3=.02*W3;
B=ID-W3;D=B*W;S=W'♦D;V=DIAG(S);S=SQRT(V);U=INV(S);Z=D♦U;
FETCH F DATA=FSCORE;SC=Z♦F;
OUTPUT SC OUT=SCOR(KEEP=COL1-COL2);
DATA SCOR;SET SCOR;
IF _N_ LE 25 THEN GP=1;IF _N_ GT 25 THEN GP=2;
PROC PLOT;PLOT COL2♦COL1=GP;

```

THE GEORGE WASHINGTON UNIVERSITY
Program in Logistics
Distribution List for Technical Papers

The George Washington University
Office of Sponsored Research
Library
Vice President H. F. Bright
Dean Harold Liebowitz
Dean Henry Solomon

ONR
Chief of Naval Research
(Codes 200, 434)
Resident Representative

OPNAV
OP-40
DCMO, Logistics
Navy Dept Library
NAVDATA Automation Cmd
OP-964

Naval Aviation Integrated Log Support

NARDAC Tech Library

Naval Electronics Lab Library

Naval Facilities Eng Cmd Tech Library

Naval Ordnance Station
Louisville, Ky.
Indian Head, Md.

Naval Ordnance Sys Cmd Library

Naval Research Branch Office
Boston
Chicago
New York
Pasadena
San Francisco

Naval Ship Eng Center
Philadelphia, Pa.
Washington, DC

Naval Ship Res & Dev Center

Naval Sea Systems Command
PMS 30611
Tech Library
Code 073

Naval Supply Systems Command
Library
Operations and Inventory Analysis

Naval War College Library
Newport

BUPERS Tech Library

FMSO

Integrated Sea Lift Study

USN Ammo Depot Earle

USN Postgrad School Monterey
Library
Dr Jack R. Forsting
Prof C. R. Jones

US Marine Corps
Commandant
Deputy Chief of Staff, R&D

Marine Corps School Quantico
Landing Force Dev Ctr
Logistics Officer
Commanding Officer
USS Francis Marion (LPA-249)

Armed Forces Industrial College

Armed Forces Staff College

Army War College Library
Carlisle Barracks

Army Cmd & Gen Staff College

Army Logistics Mgt Center
Fort Lee

Commanding Officer, USALDSMA
New Cumberland Army Depot

Army Inventory Res Ofc
Philadelphia

Air Force Headquarters
AFADS-3
LEXY
SAF/ALC

Griffiss Air Force Base
Reliability Analysis Center

Gunter Air Force Base
AFLMC/XR

Maxwell Air Force Base Library

Wright-Patterson Air Force Base
Log Command
Research Sch Log
AFALD/XR

Defense Documentation Center

National Academy of Sciences
Maritime Transportation Res Board Library

National Bureau of Standards
Dr B. H. Colvin
Dr Joan Rosenblatt

National Science Foundation

National Security Agency

Weapon Systems Evaluation Group

British Navy Staff

National Defense Hdqtrs, Ottawa
Logistics, OR Analysis Establishment

American Power Jet Co
George Chernowitz

General Dynamics, Pomona

General Research Corp
Dr Hugh Cole
Library

Logistics Management Institute
Dr Murray A. Geisler

MATHEC
Dr Eliot Feldman

Rand Corporation
Library

Carnegie-Mellon University
Dean H. A. Simon
Prof G. Thompson

Case Western Reserve University
Prof B. V. Dean
Prof M. Mesarovic
Prof S. Zacks

Cornell University
Prof R. E. Bechofer
Prof R. W. Conway
Prof Andrew Schultz, Jr.

Cowles Foundation for Research in Economics
Prof Herbert Scarf
Prof Martin Shubik

Florida State University
Prof R. A. Bradley

Harvard University
Prof K. J. Arrow
Prof W. G. Cochran
Prof Arthur Schleifer, Jr.

Princeton University
Prof A. W. Tucker
Prof J. W. Tukey
Prof Geoffrey S. Watson

Purdue University
 Prof S. S. Gupta
 Prof H. Rubin
 Prof Andrew Whinston

Stanford University
 Prof T. W. Anderson
 Prof G. B. Dantzig
 Prof F. S. Hillier
 Prof D. L. Iglehart
 Prof Samuel Karlin
 Prof C. J. Lieberman
 Prof Herbert Solomon
 Prof A. F. Veinott, Jr.

University of California, Berkeley
 Prof R. E. Barlow
 Prof D. Gale
 Prof Jack Klefer
 Prof Rosedith Sitgreaves

University of California, Los Angeles
 Prof J. R. Jackson
 Prof R. R. O'Neill

University of North Carolina
 Prof W. L. Smith
 Prof M. R. Leadbetter

University of Pennsylvania
 Prof Russell Ackoff
 Prof Thomas L. Saaty

University of Texas
 Prof A. Charnes

Yale University
 Prof F. J. Anscombe
 Prof I. R. Savage

Prof Z. W. Birnbaum
 University of Washington

Prof B. H. Bissinger
 The Pennsylvania State University

Prof Seth Bonder
 University of Michigan

Prof G. E. P. Box
 University of Wisconsin

Dr Jerome Bracken
 Institute for Defense Analyses

Prof H. Chernoff
 Mass. Institute of Technology

Prof Arthur Cohen
 Rutgers - The State University

Mr Wallace M. Cohen
 US General Accounting Office

Prof C. Derman
 Columbia University

Prof Masao Fukushima
 Kyoto University

Prof Saul I. Cass
 University of Maryland

Dr Donald P. Caver
 Carmel, California

Prof Amrit L. Goel
 Syracuse University

Prof J. F. Hannan
 Michigan State University

Prof H. O. Hartley
 Texas A & M Foundation

Mr Gerald F. Hein
 NASA, Lewis Research Center

Prof W. M. Hirsch
 Courant Institute

Dr Alan 'I. Hoffman
 IBM, Yorktown Heights

Prof John R. Isbell
 State University of New York, Amherst

Dr J. L. Jain
 University of Delhi

Prof J. H. K. Kao
 Polytech Institute of New York

Prof W. Kruskal
 University of Chicago

Mr S. Kumar
 University of Madras

Prof C. E. Lemke
 Rensselaer Polytech Institute

Prof Loynes
 University of Sheffield, England

Prof Steven Nahmias
 University of Pittsburgh

Prof D. B. Owen
 Southern Methodist University

Prof E. Parzen
 Texas A & M University

Prof H. O. Posten
 University of Connecticut

Prof R. Remage, Jr.
 University of Delaware

Prof Hans Riedwyl
 University of Bern

Dr Fred Rigby
 Texas Tech College

Mr David Rosenblatt
 Washington, D. C.

Prof M. Rosenblatt
 University of California, San Diego

Prof Alan J. Rowe
 University of Southern California

Prof A. H. Rubenstein
 Northwestern University

Dr M. E. Salvason
 West Los Angeles

Prof Edward A. Silver
 University of Waterloo, Canada

Prof M. J. Sobel
 Georgia Inst of Technology

Prof R. M. Thrall
 Rice University

Dr S. Vajda
 University of Sussex, England

Prof T. M. Whitin
 Wesleyan University

Prof Jacob Wolfowitz
 University of South Florida

Prof Max A. Woodbury
 Duke University