35975

CANADIAN THESES ON MICROFICHE

THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR/*NOM DE L'AUTEUR* DAVID W. SCOTT

TITLE OF THESIS/*TITRE DE LA THÈSE* ROTATION OF CANONICAL VARIATES USING NONIDENTICAL ORTHONORMAL TRANSFORMATIONS

UNIVERSITY/*UNIVERSITÉ* SIMON FRASER UNIVERSITY

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* MASTER OF ARTS

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1977

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* DR. RAYMOND F. KOOPMAN

DATED/*DATÉ* August 17, 1977 SIGNED/*SIGNÉ*

PERMANENT ADDRESS/*RÉSIDENCE FIXÉ* [redacted]

ROTATION OF CANONICAL VARIATES USING NONIDENTICAL ORTHONORMAL TRANSFORMATIONS

by

David William Scott

B.A. cum laude, Kalamazoo College, 1971

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Psychology

SIMON FRASER UNIVERSITY

March 1977

APPROVAL

Name: David William Scott

Degree: Master of Arts

Title of Thesis: Rotation of Canonical Variates Using Nonidentical Orthonormal Transformations

Examining Committee:

Chairman: Dr. W. Turnbull

Dr. R.F. Koopman
Senior Supervisor

Dr. C.B. Crawford

Dr. M.A. Stephens
External Examiner
Professor
Department of Mathematics
Simon Fraser University

Date Approved: 25 MAR 77

ABSTRACT

This thesis examines the theoretical and practical usefulness of rotated solutions in canonical analysis. A brief description of the theory of canonical analysis shows that a rationale can be established for rotating sets of canonical variates nonidentically, i.e., applying orthonormal transformations to both sets of canonical variates, where the two transformations are not necessarily the same. Some theory regarding a 'number of factors' analogue in canonical analysis is also discussed, in particular with respect to the effect, on a rotated solution, of any decision regarding the number of canonical correlations different from zero. Basic relations concerning rotation of canonical variates are described; it is seen that many overall scalar-valued measures of relationship between two variable sets are left unchanged when rotated variate sets are considered. A pair of rotation criteria are derived which employ a variant of the Varimax (Kaiser, 1958) criterion. The first of these, the 'raw' criterion, is analogous to 'raw' Varimax; the second, the 'normalized' criterion, is like the familiar 'normal' Varimax in that each element of the loading matrix to be rotated is scaled by its row sum of squares before rotation takes place.

Five numerical examples are presented. Two of these are drawn from the educational research literature, two are drawn

studies using personality measures, and one examines some original material involving two kinds of measures of community crime. In all cases, rotation of canonical variates is seen to produce 'cleaner' results, i.e., results more in keeping with the heuristic principles of 'simple structure'. A few of the rotations yielded striking improvements in the intuitive appeal of the results. Examination of the matrices of correlations between rotated canonical variates shows that 'normalized' rotation affects the between-set correlation structure less than does 'raw' rotation. That 'normalized' rotation leads to just as intuitively simple interpretation of results as does 'raw' rotation indicates a certain economy in rotations to this criterion.

Certain avenues for further research, especially regarding the 'number of factors' problem in canonical analysis, are indicated.

ACKNOWLEDGEMENTS

Dr. R.F. Koopman has given me considerable assistance in formulating the research problem and implementing the required computer analysis, particularly in the early stages of the project. For this, and for his remarkable patience, forbearance, and encouragement during my entire graduate school career, I thank him warmly.

Dr. C.B. Crawford originally suggested the thesis topic and suggested many worthwhile stylistic changes in the text. I benefitted from many informal discussions with him during the course of the project.

Sheila Talley typed the text of the thesis with truly amazing speed and accuracy, and tolerance for all those extra bits I kept giving her.

While he was never directly involved in my thesis research, Dr. Lorne Kendall, one of the earliest to apply canonical analysis in Psychology, showed a keen interest throughout. His untimely death on March 6, 1977 came as a shock and a loss to me. He was to have participated in my thesis examination, and he was missed.

CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Psychologists are often interested in describing the correlational relationship between two sets of variables. For example, an investigator may be interested in the relationship between intelligence and creativity, where each construct is measured by an entire battery of psychological tests. Canonical analysis, or canonical correlation, is a multivariate technique developed by Hotelling (1935, 1936) to treat this problem. Due to its computational complexity, canonical analysis lay virtually dormant for decades. Recently, with the advent of widely-available computer routines, it has assumed greater importance in the data-analytic repertoire.

Along with the renewed interest in canonical analysis, however, came the rather serious problem of interpreting its results. Many researchers find that the nature of any multivariate relationships found by canonical analysis can be disappointingly obscure. Tatsuoka (1971), while he does not deal with this problem in direct terms, alludes to it when he comforts:

> ...the dimensions of one domain (such as personality) that are strongly associated with those of another domain (such as academic achievement) are not necessarily susceptible to 'meaningful' verbal descriptions within the framework of our intuitive, everyday concepts. It may be that subsequent research will show that precisely these 'nonintuitive' dimensions represented by the canonical variates are of greater scientific import. (p. 191)

To help researchers make sense of canonical analysis, many recent methodology texts (e.g., Cooley and Lohnes, 1971; Harris, 1975; Tatsuoka, 1971; Van De Geer, 1971) have called attention to the similarities between canonical analysis and single-set techniques such as factor analysis and component analysis. In fact, Cooley and Lohnes (1971, chap. 6) go so far as to refer to 'canonical factors', and Tatsuoka (1971, pp. 183, 190) calls canonical analysis 'a double-barrelled component analysis.' To be sure, all three types of analysis belong to the family of multivariate techniques which attempt to summarize the relevant information in a (large) set of observed variables with a (small) set of hypothetical variables which form an orthonormal basis for a subspace of the original variable space. This basis is positioned arbitrarily, however, and a user of component analysis or factor analysis typically rotates an initial component or factor solution to a position where the basis vectors are easier to interpret in terms of observed variables. A user of canonical analysis, however, is generally left to interpret an unrotated solution.

Cliff and Krus (1976) have presented a canonical analysis in which they transform, or rotate, the two sets of canonical variate vectors to some advantage in interpreting their results. They propose, and use, a single orthonormal transformation which rotates each set of variates identically. This paper describes a rationale for considering nonidentically rotated canonical variate sets, and adapts a derivative of the Varimax method (Kaiser, 1958) to the concerns of canonical analysis as a means of obtaining appropriate transformations. It is shown that nonidentical rotations leave unchanged many important measures of the overall correlational similarity between two sets of variables. Also, rotations performed on actual examples of canonical analysis from the behavioural science literature demonstrate that canonical analysis can in fact make good intuitive sense when nonidentical transformations are admitted.

2. SOME THEORY FROM THE STATISTICAL LITERATURE

2.1 What is Canonical Analysis?

Canonical analysis is a generalization of bivariate linear correlation developed by Hotelling (1935, 1936). In this technique, a linear combination, or weighted sum, of each of two collections of observed variables is obtained: the weights used to determine these 'canonical variates' are chosen so that the bivariate linear correlation, or 'canonical correlation', between the two combinations is a maximum. The process can continue: A further linear combination can be obtained from each variable set, subject to the same maximum inter-set correlation constraint, but with the further restriction that each new variate is orthogonal to the first two derived variates. As many pairs of canonical variates as there are observed variables in the smaller of the two sets can be obtained in this way; each successively-derived variate is orthogonal to all previously-derived variates, and correlates maximally with its 'sister' variate in the other set.

The theory underlying canonical analysis is well described in many texts intended for practicing researchers, notably those by Bock (1975), Cooley and Lohnes (1971), Harris (1975), Morrison (1976), Press (1972), Tatsuoka (1971), and Van De Geer (1971). More mathematical treatments of the subject can be found in Anderson (1958), Dempster (1969), Kendall (1961), and Rao (1965), as well as in a primary reference by Hotelling (1936).

This chapter presents some of the basic equations of the technique, to establish notation and clarify the nature of the quantities with which the rest of the paper deals. It also considers an interesting analogue, due to Rao (1965), of the 'number of factors' problem in factor analysis, and motivates the discussion of the rotation of canonical variates with a geometric representation of canonical analysis, much of which is drawn from Dempster (1969).

2.2 Basic equations

The raw data for canonical analysis is a partitioned matrix $X:N\times n = [X_1,X_2]$ where $X_1:N\times n_1$ and $X_2:N\times n_2$ contain the data for two distinct sets of $n_1+n_2 = n$ variables. Following Morrison (1976), this discussion assumes that the N data (row) vectors constituting X have been drawn from an n-dimensional population with covariance matrix $\Sigma = \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}$, where $\Sigma_{21} = \Sigma_{12}'$. For convenience, and without losing generality, we can also assume that the data is mean-corrected, that is, that all column sums of X are zero, and that $n_1 \le n_2$. The sample covariance matrix of X is then given by

$$C:n\times n = \frac{1}{N} X'X = \frac{1}{N}\begin{bmatrix}X_1'\\ X_2'\end{bmatrix}[X_1,X_2] = \begin{bmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{bmatrix}$$

where C is partitioned the same way as Σ, $C_{11}:n_1\times n_1$ and $C_{22}:n_2\times n_2$ are symmetric within-set covariance matrices, and $C_{12}:n_1\times n_2 = C_{21}'$ is a between-set covariance matrix. We further assume that C is positive definite. The initial approach to relating X_1 and X_2 involves coefficients $w_1:n_1\times 1$ and $w_2:n_2\times 1$ such that the linear combinations $y_1 = X_1w_1$ and $y_2 = X_2w_2$ are maximally correlated. The estimate of the covariance matrix of these linear compounds is

$$\begin{bmatrix} w_1' \\ w_2' \end{bmatrix} C[w_1,w_2] = \begin{bmatrix} w_1' C_{11}w_1 & w_1' C_{12}w_2 \\ w_2' C_{21}w_1 & w_2' C_{22}w_2 \end{bmatrix}$$

(note that $w_1'C_{12}w_2 = w_2'C_{21}w_1$), and the estimate of their squared correlation is

$$r_1^2 = \frac{(w_1'C_{12}w_2)^2}{(w_1'C_{11}w_1)(w_2'C_{22}w_2)}$$

We wish to determine values of w_1 and w_2 which maximize r_1^2 .

The solution to this problem involves solving a system of equations in w_1 and w_2 . To obtain a unique solution, we impose the following constraint on the scales of $y_1 = X_1w_1$ and $y_2 = X_2w_2$:

$$w_1'C_{11}w_1 = w_2'C_{22}w_2 = 1 \quad .$$

Using Lagrange multipliers μ_1 and μ_2 we can incorporate this constraint into a function to be minimized by writing

$$f(w_1,w_2) = (w_1'C_{12}w_2)^2 - \mu_1(w_1'C_{11}w_1-1) - \mu_2(w_2'C_{22}w_2-1)$$

Taking derivatives of f with respect to w_1 and w_2 , and setting these derivatives equal to zero, we obtain the following two matrix equations:

$$(1)\qquad \begin{aligned} -\mu_1 C_{11} w_1 + (w_1' C_{12} \mu_2)\; C_{12} w_2 &= \underline{0}_1 \\ (w_1' C_{12} w_2)\; C_{21} w_1 - \mu_2 C_{22} w_2 &= \underline{0}_2 \end{aligned}$$

where $\underline{0}_1$ and $\underline{0}_2$ are zero vectors of appropriate dimensionality. Premultiplying the first equation of (1) by w_1' and the second by w_2' , yields

7

$$-\mu_1 (w_1' C_{11} w_1) + (w_1' C_{12} w_2)^2 = \underline{0}_1$$

$$(w_1' C_{12} w_2)^2 - \mu_2 (w_2' C_{22} w_2) = \underline{0}_2 \; .$$

Solving for μ_1 in the first equation and μ_2 in the second yields

$$\mu_1 = \mu_2 = (w_1' C_{12} w_2)^2$$

If we substitute $\lambda^2 = \mu_1 = \mu_2$,

the matrix equations (1) can be written

$$-\lambda^2 C_{11} w_1 + \lambda \cancel{C_{12} w_2} = \underline{0}_1$$

$$\lambda C_{21} w_1 - \lambda^2 C_{22} w_2 = \underline{0}_2$$

or

$$(2)\qquad M \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \underline{0}_1 \\ \underline{0}_2 \end{bmatrix}$$

$$M = \begin{bmatrix} -\lambda^2 C_{11} & \lambda C_{12} \\ \lambda C_{21} & -\lambda C_{22} \end{bmatrix}$$

If $C_{12} = \{0\}$ then $w_1' C_{12} w_2 = 0$ for any choice of w_1 and w_2 .

However, (2) has a nontrivial solution only if $|M| = 0$. Results concerning the determinant of a partitioned matrix (see, for example, Morrison, 1976, pp. 67-68), as well as the fact that since C is positive definite, C_{11} and C_{22} must also be, yield that

$$|M| = |-\lambda^2 C_{22}|\,|-\lambda^2 C_{11} - \lambda^2 C_{12}(-\lambda^2 C_{22})^{-1} C_{21}| = 0$$

$$= |-\lambda^2 C_{11}|\,|-\lambda^2 C_{22} - \lambda^2 D_{21}(-\lambda^2 C_{11})^{-1} C_{12}| = 0 \quad .$$

Simplifying,

$$[-\lambda^2 C_{11} + C_{12}C_{22}^{-1}C_{21}] = 0$$

$$[-\lambda^2 C_{22} + C_{21}C_{11}^{-1}C_{12}] = 0$$

or

$$[C_{11}^{-1}C_{12}C_{22}^{-1}C_{21} - \lambda^2 I] = 0$$

$$[C_{22}^{-1}C_{21}C_{11}^{-1}C_{12} - \lambda^2 I] = 0 \quad .$$

The largest eigenvalue of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$, or equivalently of $C_{22}^{-1}C_{21}C_{11}^{-1}C_{12}$, as these two matrices are cyclic permutations of one another, is the maximum value of r_1^2 over all values of w_1 and w_2. Moreover, Anderson (1958) shows that the coefficient vector w_1 is the appropriately-scaled eigenvector associated with the largest eigenvalue of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$, and that the coefficient vector w_2 is given by

$$(3) \qquad w_2 = \frac{1}{\lambda} C_{22}^{-1}C_{21}w_1 \quad .$$

Thus we have found the largest canonical correlation and the weight vectors, or canonical coefficients, used to form its associated canonical variates.

If $n_1 \leq n_2$ is greater than 1, however, it becomes of interest to consider any further combinations of X_1 and X_2 which are maximally correlated, subject to the restriction that each be uncorrelated with the first canonical variates y_1 and y_2 . If

$$y_1^{(1)} = y_1 = X_1 w_1^{(1)}$$

and

$$y_2^{(1)} = y_2 = X_2 w_2^{(1)}$$

denote these first canonical variates (and their associated canonical coefficients), the new linear combinations can be written

$$y_1^{(2)} = X_1 w_1^{(2)}$$

and

$$y_2^{(2)} = X_2 w_2^{(2)}$$

where

$$y_1^{(2)'} y_1^{(1)} = y_1^{(2)'} y_2^{(1)} = y_2^{(2)'} y_1^{(1)} = y_2^{(2)'} y_2^{(1)} = 0 \quad ,$$

$$\frac{1}{N}\, y_1^{(2)'} y_2^{(2)} = r_2$$

is another canonical correlation, and $y_1^{(2)}$ and $y_2^{(2)}$ are its associated canonical variates. Once again, to obtain a unique solution, we choose $w_1^{(2)}$ and $w_2^{(2)}$ so that

$\text{var}(y_1^{(2)}) = \text{var}(y_2^{(2)}) = 1$. If X_1 and X_2 are each composed of linearly independent columns, it would seem possible to obtain n_1 orthogonal variates from X_1 , each maximally correlated with a similar variate obtained from X_2 -- in other words, to extract $2n_1$, orthogonal canonical variates $y_1^{(i)}$ and $y_2^{(i)}$, $i = 1,\ldots,n_1$, such that

$$y_1^{(i)'} y_1^{(j)} = y_1^{(i)'} y_2^{(j)} = y_2^{(i)'} y_2^{(j)} = 0 , \quad i \neq j ,$$

and where the $r_i = \frac{1}{N} y_1^{(i)'} y_2^{(i)}$ constitute the set of canonical correlations between X_1 and X_2 . Anderson (1958) gives detailed proof that it is in fact possible to do so, and that this is equivalent to finding the n_1 nonzero eigenvalues of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$ and a set of associated eigenvectors.
As such,

$$C_{11}^{-1}C_{12}C_{22}^{-1}C_{21} = W_1\Lambda^2W_1^{-1}$$

where $W_1 : n_1 \times n_1$ is the matrix of coefficients for combining the X_1 variables to form the $y_1^{(i)}$, and Λ^2 is a diagonal matrix of squared canonical correlations between X_1 and X_2 . If we adopt the convention

$$\lambda_1^2 > \lambda_2^2 > \ldots > \lambda_n^2 ,$$

then $y_1^{(i)} = X_1 w_1^{(i)}$, where $w_1^{(i)}$ is now the i^{th} column of W_1 . (3) can be generalized to solve for the matrix W_2 :

$$W_2 = C_{22}^{-1}C_{21}W_1\Lambda^{-1} ,$$

then

$$y_2^{(i)} = X_2 w_2^{(i)}$$

and

$$\frac{1}{N}\, y_1^{(i)'} y_2^{(i)} = r_i = \lambda_i \quad .$$

In summary,

$$Y_1 = X_1 W_1$$

$$Y_2 = X_2 W_2$$

where the respective i^{th} columns of Y_1 and Y_2 are $y_1^{(i)}$ and $y_2^{(i)}$, the i^{th} pair of canonical variates. Also,

$$\frac{1}{N}\, Y_1' Y_2 = \Lambda \quad , \tag{4}$$

the diagonal matrix of canonical correlations, ordered largest to smallest. The λ_i are the square roots of the n_1 nonzero eigenvalues of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$. Further,

$$\frac{1}{N}\, Y_1' Y_1 = I$$

$$\frac{1}{N}\, Y_2' Y_2 = I$$

that is, the canonical variates are uncorrelated with one another within-sets.

2.3 The 'number of factors' problem in canonical analysis

Rao (1965) has presented an interesting conceptualization of canonical analysis which provides some insight for the common finding that not all of the canonical variates obtained using the method of section 2.2 are actually meaningful in describing between-set relationships. This situation finds its single-set analogue in the classical 'number of factors' problem in factor analysis (Harman, 1966; Lawley and Maxwell, 1971; Mulaik, 1972). This section sketches Rao's argument.

Rao suggests that a model for canonical analysis can be drawn up as follows:

$$X_1 = FP_1' + S_1$$

$$X_2 = FP_2' + S_2$$

where $F:N\times m$ is a matrix of scores on a hypothetical, unobserved set of m 'common factors' which contribute to both X_1 and X_2 ; $S_1:N\times n_1$ and $S_2:N\times n_2$ are matrices of scores on two hypothetical unobserved sets of 'specific factors' which contribute to either X_1 or S_2 , but not both; and $P_1:n_1\times m$ and $P_2:n_2\times m$ are matrices of 'common-factor' loadings. F , S_1 , and S_2 are here assumed to be column mean-corrected. Rao includes the following restrictions in his definition of the model:

$$\frac{1}{N} F'S_1 = \frac{1}{N} F'S_2 = \frac{1}{N} S_1'S_2 = 0$$

and

$$\frac{1}{N} F'F = I$$

that is, the 'common factors' are uncorrelated with one another and with the 'specific factors', and factors 'specific' to X_1 are uncorrelated with those 'specific' to X_2 .

Note that the use of 'common' and 'specific' here differs from the traditional use of these terms in the context of factor analysis. With this in mind, the quotation marks around these terms can be omitted.

The partitioned matrix C is redefined:

$$C_{11} = \frac{1}{N}(FP_1'+S_1)'(FP_1'+S_1) =$$

$$= \frac{1}{N}[P_1F'FP_1'+P_1F'S_1+S_1'FP_1'+S_1'S_1]$$

$$= \frac{1}{N}[P_1P_1'+S_1'S_1]$$

$$= \frac{1}{N}P_1P_1'+D_1$$

where D_1 is the covariance matrix of S_1 . Similarly,

$$C_{22} = \frac{1}{N}P_2P_2'+D_2$$

where D_2 is the covariance matrix of S_2 . On the other hand,

$$C_{12} = \frac{1}{N}(FP_1'+S_1)'(FP_2'+S_2) =$$

$$= \frac{1}{N}[P_1F'FP_2'+P_1F'S_2+S_1'FP_2'+S_1'S_2]$$

$$= \frac{1}{N}P_1P_2'$$

This relation shows that any association between X_1 and X_2 can be attributed solely to common factors, while within-set relationships involve contributions from common and specific factors.

It now becomes of interest to determine the number of common factors, that is, the number of columns of F. Rao equates a quantity called the 'effective number of common factors' with the rank of Σ_{12}, the population analogue of C_{12}. Not surprisingly, this quantity is equal to the number of nonzero eigenvalues of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Since we necessarily deal with sample estimates of these population matrices, however, it will happen that the rank of C_{12} is equal to n_1 ($\leq n_2$, by convention), even when the rank of Σ_{12} is equal to a smaller value. Since many of the eigenvalues of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$ may be close to zero, however, one approach to determining the effective number of common factors is to test the hypothesis that these eigenvalues are in fact identically zero. An approximate test of this hypothesis, due to Bartlett (1941, 1947), is discussed in section 2.5.

Rao suggests that, under ideal circumstances, an even smaller number of 'dominant common factors' will account for the nontrivial information about the relationship between X_1 and X_2. He allows, however, that since the inferential machinery to estimate the number of these dominant common factors does not exist, less mathematically rigorous methods, based on the relative magnitudes of the eigenvalues of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$, may be appropriate. The Bartlett method of section 2.5 provides one way to settle on a value of m, the number of common factors, using an approximate statistical test. Notably, though, studies into the most acceptable number of common factors in canonical analysis are not available, and the problem is far from solved.

Common factors are typically not estimated in canonical analysis. As a later section of this paper discusses, however, some results reported by Carroll (1968), in the context of generalizing canonical analysis to more than two sets of observed variables, bear on the notion of common factors.

2.4 A geometric representation of canonical analysis

Dempster (1969) provides a geometric description of canonical analysis in which the technique becomes a means of describing the system of angles between particular orthonormal bases of two subspaces of the variable space defined by the columns of X . Let U denote the space spanned by all columns of X . Then U_1 and U_2 are two complementary subspaces of U: U_1 the space spanned by X_1 , and U_2 the space spanned by X_2 . The bases $Y_1 = X_1W_1$ and $Y_2 = X_2W_2$ form sets of orthogonal basis vectors for U_1 and U_2 , respectively, with the property that each vector in Y_1 is a 'best linear predictor' of a vector in Y_2 , and vice versa. Note that Y_2 actually spans an n_1-dimensional subspace of U_2 : call this subspace $U_{2(1)}$. The $(n_2 - n_1)$ remaining dimensions of U_2 are orthogonal to U_1 , in that no linear predictor can be formed in U_1 which will make an angle different from $\Pi/2$ with any vector formed in this 'leftover' vector space. The cosines of the angles between the vectors in

Y_1 and those in Y_2, which become the elements of the matrix Λ (from (4)) are the nonzero canonical correlations on the diagonal of Λ and zero elsewhere. These canonical correlations are interpreted as measures of similarity between pairs of basis vectors in U_1 and $U_{2(1)}$, and hence, jointly, as a measure of overall similarity between the two subspaces of U.

The dimensionality, n_1, of both U_1 and $U_{2(1)}$ is equated with the number of nonzero canonical correlations, which is also the rank of C_{12}. Note that since U_1 and $U_{2(1)}$ are linearly independent, although correlated, we need a space of at least $2n_1$ dimensions to represent both subspaces simultaneously. Also, m becomes the dimensionality of even smaller subspaces U_1^* and $U_{2(1)}^*$ which are spanned by basis vectors summarizing all the nontrivial information about the relationship between X_1 and X_2.

2.5 Estimation and significance testing

If X is assumed drawn from an n-dimensional multivariate normal population with arbitrary mean vector μ and covariance matrix Σ, Anderson (1958) shows that the method of section 2.2 leads to maximum likelihood estimates of the population canonical correlations and canonical variates. Moreover, the largest eigenvalue of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$ (the largest squared canonical correlation) follows the greatest characteristic root distribution and can be used to test the hypothesis that all such eigenvalues are zero, that is, there is no relationship between x_1 and x_2 ..

As the results of sections 2.3 have shown, however, the number of nonzero eigenvalues of $C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$ will give an idea of the number of statistically important dimensions by which X_1 and X_2 are related. Bartlett (1941, 1947) has proposed an approximate, large-sample method for testing the simultaneous difference of a subset of these eigenvalues from zero, based on a distribution which is, asymptotically, an approximate chi-square. Bartlett's technique is used popularly in applications of canonical analysis. Harris (1975) argues on logical grounds that this technique is never appropriate, and suggests using the greatest characteristic root test, with a modified degree of freedom parameter, to test the significance of all the λ_i^2 . He mentions, however, that such tests may be conservative for all but the largest λ_i^2 and offers no sampling-theory justification for using such a method.

3. ROTATION OF CANONICAL VARIATES

3.1 Basic relations

This chapter treats the problem of rotating canonical variates to improve their interpretability. Our task will be to find a canonical-analysis analogue of 'simple structure' (Thurstone, 1947) in matrices of correlations between the bases Y_1 and Y_2, and the observations X_1 and X_2. We assume that the value of m has been determined, and redefine Y_1 and Y_2 to be of dimensions $N \times m$ and $N \times m$, respectively.

Consider two orthonormal matrices $T_1 : m \times m$ and $T_2 : m \times m$, where T_1 and T_2 are not necessarily the same. The expression

$$Z_1 = Y_1 T_1$$

represents the orthogonal rotation of the m basis vectors of U_1^*.

$$Z_2 = Y_2 T_2$$

similarly represents an orthogonal rotation of the basis vectors of $U_{2(1)}^*$. Z_1 and Z_2 are then alternative bases for these subspaces. Put another way, Z_1 and Z_2 are sets of rotated canonical variates.

Since $Y_1 = X_1 W_1$ and $Y_2 = X_2 W_2$, we observe that

$$Z_1 = Y_1 T_1 = X_1 W_1 T_1 = X_1 V_1$$

and
$$Z_2 = X_2 V_2$$

where $V_1 = W_1T_1$ and $V_2 = W_2T_2$ are matrices of coefficients for obtaining Z_1 and Z_2 from X_1 and X_2. Moreover, if

$$A_1 = \frac{1}{N}(X_1'Y_1)$$

and

$$A_2 = \frac{1}{N}(X_2'Y_2)$$

then

$$B_1 = \frac{1}{N}(X_1'Z_1) = \frac{1}{N}(X_1'Y_1T_1) = A_1T_1$$

and

$$B_2 = A_2T_2$$

are matrices of structure correlations between Z_1 and Z_2 and the observed matrices from which these variates are derived. The transformations T_1 and T_2 can thus be applied to either the coefficient matrices W_1 and W_2 or directly to the structure matrices A_1 and A_2 in obtaining results from the rotated solution.

3.2 The effect of rotation on measures of between-set relationship

Importantly, various scalar-valued measures of the relationship between X_1 and X_2 remain unchanged under the orthogonal rotations T_1 and T_2. Originally, Hotelling (1936; see also Anderson, 1958, pp. 244-45) proposed a 'vector coefficient of alienation'

$$(5)\qquad \frac{|C_{11}-C_{12}C_{22}^{-1}C_{21}|}{|C|} = \frac{|C|}{|C_{11}||C_{22}|}$$

and a corresponding 'vector correlation coefficient'

$$\text{(6)} \qquad \frac{|C_{12}C_{22}^{-1}C_{21}|}{|C_{11}|} ,$$

based on the notion of 'generalized variance' (Wilks, 1932), as multivariate analogues of the coefficients of alienation and (simple or multiple) correlation in regression situations with a single dependent variable. Since Hotelling's measures depend only on C and its submatrices, their values and interpretation will not change whether the canonical variates, which are merely linear combinations of X_1 and X_2, are obtained using coefficient matrices W_1 and W_2 or V_1 and V_2.

Note that (5), which is equal to $\prod_{i=1}^{m} (1-r_i^2)$, and (6), which is equal to $\prod_{i=1}^{m} r_i^2$, suffer the disadvantage that they can each go to zero based on the value of only a single (typically nonrepresentative) canonical correlation. In particular, (5) is near-zero if any of the canonical correlations is near-perfect. A more likely result, however, is that one or more of the canonical correlations is near-zero. In such a case, (6) becomes very small, even if some of the canonical correlations are very large.

Another type of scalar-valued measure is found in various additive functions of the squared correlations between the basis vectors of one set with either the basis vectors or the original variables of the other set. One such function is simply the sum of the squared canonical correlations, $\sum_{k=1}^{m} r_k^2$, or in matrix notation, $tr\ \Lambda^2$. But consider the matrix of correlations between rotated variates Z_1 and Z_2 :

$$(7) \qquad L = \frac{1}{N} Z_1'Z_2 = \frac{1}{N} T_1'Y_1'Y_2T_2 = T_1'\Lambda T_2$$

This matrix contains, in its i^{th} row, the correlations between the rotated variate $z_1^{(i)}$ and all the variates Z_2 . Since the canonical variates are uncorrelated within sets, the sum of squares in the i^{th} row of L is the squared multiple correlation between $z_1^{(i)}$ and $z_2^{(j)}$, $j = 1,\ldots,m$ . . In matrix notation, the diagonal of

$$S = LL'$$

will contain all m of these squared multiple correlations. But

$$S = LL' = T_1'\Lambda T_2T_2'\Lambda T_1 = T_1'\Lambda^2T_1$$

Clearly, the eigenvalues of S are the r_k^2 , so that

$$trS = tr\Lambda^2 = \sum_{k=1}^{m} r_k^2 \ .$$

In other words, the sum of the individual squared multiple correlations between each rotated canonical variate in Z_1 and all the rotated canonical variates in Z_2 is simply the sum of the squared canonical correlations.

This relationship holds if we interchange the roles of Z_1 and Z_2 . To see this, note that the j^{th} diagonal element of

$$S^* = L'L$$

is the squared multiple correlation of $z_2^{(j)}$ with the $z_1^{(i)}$.
But

$$\text{tr } S = \text{tr } S^* = \text{tr } \Lambda^2$$

Thus tr S and tr S^* are equal to one another and to tr Λ^2 and as such are invariant under choice of orthonormal T_1 and T_2. (In the special case $T_1 = T_2 = I$, then $S = S^* = \Lambda^2$.)

A measure which is similar to tr Λ^2 in many respects is the 'redundancy index' (Stewart and Love, 1968; Miller and Farr, 1971). The redundancy of X_1 in X_2, or $\bar{R}^2_{2.1}$, is defined, using present notation, as

$$\bar{R}^2_{2.1} = \frac{1}{n_2} \sum_{k=1}^{m} r_k^2 \sum_{j=1}^{n_2} a_{2jk}^2$$

The redundancy of X_2 in X_1 is defined interchanging the variable-set subscripts 1 and 2. Miller (1975) shows that $\bar{R}^2_{2.1}$ is simply the average, over variables, of the squared multiple correlations between the m columns in Y_1 and the n_2 columns of observations in X_2. In doing so, however, he also shows that $\bar{R}^2_{2.1}$ is equal to

$$\frac{1}{n_2} \text{tr } R_{21}R_{11}^{-1}R_{12}$$

where R_{11}, R_{12}, and R_{21} are C_{11}, C_{12}, and C_{21} rescaled to correlation matrices. Ultimately, then, the redundancy index, like Hotelling's index, does not depend on the orientation of the bases of U_1^* and $U_{2(1)}^*$, that is, whether canonical variates are rotated or not.

3.3 Identical vs. nonidentical transformations

Cliff and Kruss (1976) have proposed rotating canonical variates using a single orthonormal transformation T, which is applied to Y_1 and Y_2 simultaneously. They argue, as this paper does, that certain properties of the canonical solution do not change under such a transformation, but do not offer a rationale for choosing the transformation they eventually perform. The major difference between the Cliff and Krus approach and the one presented here concerns the matrix

$$L = T_1' \Lambda T_2 \quad ,$$

L is not symmetric; an L^* defined as

$$L^* = T' \Lambda T$$

as Cliff and Krus have proposed, would be. But this symmetry implies that the i^{th} row sum of squares of L^*, which is the squared multiple correlation between $z_1^{(i)}$ and all the columns of Z_2, is equal to the i^{th} column sum of squares, which is the squared multiple correlation between $z_2^{(i)}$ and all the columns of Z_1. There is no a priori reason why this should be so, nor can any apparent interpretive utility be derived from this relation.

In fact, the requirement that the transformation matrices be identical may place an unnecessary burden on our search for some kind of 'simple structure' in B_1 and B_2 simultaneously. Each variable set will have a unique canonical variate structure, and the initial canonical solution may yield a more ideally placed basis in one subspace or the other. Hence, it will be advantageous to be able to rotate B_1 and B_2 uniquely.

3.4 The rotation criterion

The problem thus becomes one of finding suitable orthonormal transformations. One starting point is provided by Hakstian (1976) who demonstrates simultaneous rotation of two loading matrices which are assumed to represent the same number of factors, using a two-matrix extension of the general Orthomax rotation criterion (Harris and Kaiser, 1964). Hakstian does not apply his technique to canonical analysis in which two or more loading matrices are considered at the same time. Like Cliff and Krus, he seeks a single transformation matrix, but one which will maximize the following function:

$$(8) \qquad f = \sum_{k=1}^{m} \left[\frac{1}{n_1} \sum_{j=1}^{n_1} b_{1jk}^4 + \frac{1}{n_2} \sum_{j=1}^{n_2} b_{2jk}^4 \right]$$

$$-w \sum_{k=1}^{m} \left[\frac{1}{n_1^2} \left(\sum_{j=1}^{n_1} b_{1jk}^2 \right)^2 + \frac{1}{n_2^2} \left(\Sigma\, b_{2jk}^2 \right)^2 \right]$$

where b_{ijk} is the j,k^{th} element of the i^{th} rotated loading matrix, $i = 1,2$; m is the number of factors (a value common to both sets); n_i is the number of variables in the i^{th} loading matrix; and w is the weighting factor which, when considering a single matrix, determines the special case of the Orthomax criterion. $w = 1$ yields a generalization of Varimax (Kaiser, 1958). If we make this substitution, (8) can be rewritten:

$$(9) \qquad f = \sum_{k=1}^{m} \left\{ \frac{1}{n_1} \left[\sum_{j=1}^{n_1} b_{1jk}^4 - \frac{1}{n_1} \left(\sum_{j=1}^{n_1} b_{1jk} \right)^2 \right] + \frac{1}{n_2} \left[\sum_{j=1}^{n_2} b_{2jk}^4 - \frac{1}{n_2} \left(\sum_{j=1}^{n_2} b_{2jk}^2 \right)^2 \right] \right\}$$

This criterion seeks to maximize the Varimax criterion computed within matrices, then pooled between matrices. The basic advantage of this method lies in its search for a simple structure solution within two matrices simultaneously. An alternative method, which would apply to Varimax criterion to the supermatrix

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

could conceivably lead to small loadings in one matrix and large loadings in the other. This would clearly be an unacceptable solution.

If we relax the requirement that both matrices B_1 and B_2 be transformed identically, (9) declines in usefulness, since a Varimax rotation performed on each matrix separately would achieve the same result. At this point, however, we can recognize the emphasis in canonical analysis on the correspondence between certain canonical variates, namely the variate pairs $z_1^{(i)}$ and $z_2^{(i)}$, $i = 1,\ldots,m$. In loose language, we can try to make a pair of variates 'predict one another' by keeping the diagonal elements of L (from (7)) large, and hence the off-diagonals of L small. Canonical variates can then be interpreted not only in terms of structure correlations with the observed variables from which each was obtained, but also in terms of a representative variate 'from the other set.' The ideal solution, in which some semblance of simple structure is achieved in both B_1 and B_2 , and yet L is still dominated by its diagonal elements, would permit particularly meaningful choices of subsets of the observed variables with which to describe interbattery relationships.

With this in mind, this paper proposes weighting the k^{th} term of the outer sum in (9) by l_{kk}^2 , the k^{th} squared diagonal element of L . The function to be maximized then becomes

$$(10)\quad F = \sum_{k=1}^{m} l_{kk}^2 \left\{ \frac{1}{n_1}\left[\sum_{j=1}^{n_1} b_{1jk}^4 - \frac{1}{n_1}(\Sigma\, b_{1jk}^2)^2\right] + \frac{1}{n_2}\left[\sum_{j=1}^{n_2} b_{2jk}^4 - \frac{1}{n_2}(\sum_{j=1}^{n_2} b_{2jk}^2)^2\right] \right\}$$

A variant of this criterion is an analogue of 'normalized' Varimax (Kaiser, 1958). In the normalized case, the function to be maximized is

$$(11) \qquad F_n = \sum_{k=1}^{m} l_{kk}^2 \left\{ \frac{1}{n_1} \left[\sum_{j=1}^{n_1} \left(\frac{b_{1jk}}{h_{1j}}\right)^4 - \frac{1}{n_1} \left(\sum_{j=1}^{n_1} \frac{b_{1jk}^2}{h_{1j}^2} \right)^2 \right] + \frac{1}{n_2} \left[\sum_{j=1}^{n_2} \left(\frac{b_{2jk}}{h_{2j}}\right)^4 - \frac{1}{n_2} \left(\sum_{j=1}^{n_2} \frac{b_{2jk}^2}{h_{2j}^2} \right)^2 \right] \right\}$$

where

$$h_{1j}^2 = \sum_{k=1}^{m} b_{1jk}^2 \,, \quad j = 1,\ldots,n_1$$

and

$$h_{2j}^2 = \sum_{k=1}^{m} b_{2jk}^2 \,, \quad j = 1,\ldots,n_2 \,.$$

3.5 Implementation of the rotation procedure

Standard treatments of analytic rotation (e.g., Kaiser, 1958) have taken advantage of the fact that in a single plane, say the plane defined by factors (or components) r and s, elements of the rotated matrix are related to elements of the unrotated matrix in the following way:

$$b_{jr} = a_{jr} \cos\theta + a_{js} \sin\theta$$

$$b_{js} = -a_{jr} \sin\theta + a_{js} \cos\theta$$

where θ is the angle of rotation in the r, s plane. By treating the individual loadings in the rotated matrix as linear combinations of loadings in the unrotated matrix, with functions of θ as coefficients, the rotation criterion can be rewritten as an expression involving θ. The single-plane rotation problem is usually solved by taking the derivative of the rewritten function with respect to θ, setting this derivative equal to zero, and solving for θ.

In the present situation, such a procedure proves to be unweildy. Since the transformations T_1 and T_2 are not assumed to be identical, two angles of rotation -- call them θ_1 and θ_2 -- enter (10) in the single plane. Moreover, the presence of the l_{kk}^2 further complicates (10) even though it is possible in principle to express the post-rotation l_{kk}^2 as linear functions of the pre-rotation λ_k^2. To simplify the maximization process, a simple trial-and-error search procedure is used wherein for each $(\theta_{1j}, \theta_{2j})$, i.e., for the j^{th} pair of variates in z_1 and the pair corresponding to it in z_2, the $(\theta_{1j}, \theta_{2j})$ plane is searched until the point is found at which (10) is a maximum.

In cases where there are more than two columns each in z_1 and z_2, the same iterative procedure is followed as is customary with other applications of analytic rotation. One complete iteration is finished when the search algorithm, working plane-by-plane, has determined an initial approximation to the ultimate rotation in all possible planes (in this case, pairs of planes). Iteration continues until all changes in orientation of the rotated variates are smaller than some prespecified value, i.e., until the process converges.

4. NUMERICAL EXAMPLES

4.1 A note on the presentation of results

This chapter presents five examples of canonical analysis, four from published studies in Psychology or Education, and one involving some unpublished data. For each study, results are considered from the unrotated canonical solution and two rotated canonical solutions -- a 'raw' rotation to the criterion of (11) in section 3.4, and a 'normalized' rotation to the criterion of (12). The appendices contain three matrices for each solution: two loading matrices containing the structure correlations between the observed variable and the (rotated or unrotated) canonical variates, and matrices of between-set canonical variate intercorrelations. In an unrotated case, this latter matrix will be simply a collection of canonical correlations. Also, the loading matrices are represented by a series of two-dimensional plots, in which two orthogonal canonical variates are drawn as co-ordinate axes, and orthogonal projections of observed variables into the plane spanned by these two variates are drawn as vectors (arrows). In the plots, projections of observed variables are drawn as light-lined arrows, and are marked by arabic numerals; the variables can be identified by reference to whatever matrix is being plotted. Co-ordinate axes, and projections of canonical variates 'from the other set', are drawn as heavy-lined arrows, and are marked by roman numerals followed by a (1) or (2). A (1) means that the marked variate arises from variable set 1; a (2) indicates that it arises from set 2.

This chapter describes canonical variates in terms of these abbreviations. For example, 'variate II(1)' will mean 'canonical variate II from set 1', and so on. Note that variate II(1) is the same as $z_1^{(2)}$ in the notation from previous chapters. This change will make it easier to identify variates in the plots, and is consistent with usual factor-analytic practice of labelling plotted factors with roman numerals.

All loading matrices present canonical loadings rounded to two decimal places. Loadings greater than .30 in magnitude are marked by a (+), and those greater than .70 by a (++). Similarly, loadings less than -.30 are marked by (-), and those less than -.70 by (--). It will often be advantageous to pay particularly close attention to 'double-plus' or 'double-minus' loadings.

Most of the studies discussed here have used the Bartlett chi-square approximation in choosing m, the number of 'significant' canonical correlations. This chapter presents some evidence that doing so may lead to 'over-factoring', that is, keeping more dimensions than are actually of scientific interest.

4.2 <u>Ability tests vs. achievement tests</u> (see Appendix A)

Lohnes and Gray (1972) report a study in which they compare a component analysis and a canonical analysis of some data from the United States Office of Education Co-operative Reading Studies (Dykstra, 1968). One set of variables consists of eight tests used to assess reading readiness; the other consists of a battery of Stanford Achievement Tests.

Lohnes and Gray first treat both sets of variables as a single matrix and run a component analysis of this larger variable set. They report a substantial 'general factor' on which all tests, ability and achievement, load very highly, and a smaller second component (see Table A-I). They go on to use the results of a canonical analysis between the ability and achievement tests to reaffirm the existence of this 'general factor'.

Close scrutiny of their component analysis results, however, reveals that Lohnes and Gray report a loading matrix obtained from unrotated components. Table A-II shows the Lohnes and Gray loading matrix after if has been rotated to the normalized Varimax criterion. As might be expected, the rotated solution shows components substantially related to 'ability' and 'achievement'. Note, however, that the largest loadings of ability tests on the 'achievement' component (I) are from the Murphy-Durrell Phonemes and Letter Names tests; the largest loading of an achievement test on the 'ability' component (II) is from the Stanford Achievement Test in Vocabulary.

The results of the Lohnes and Gray canonical analysis are reproduced in Table A-III, and plots of these components are shown in Figure A-1. The 'general factor' interpretation is apparent. The results of a raw rotation of the canonical variates are presented in Tables A-IV and A-V; the associated plot is Figure A-2. The rotated variate sets show an interesting contrast to the unrotated variates. For one thing, the rotation procedure places each variable vector in the first quadrant of whatever variate plane it is projected into. While much of the 'general-factor' appearance of the analysis is still present, the 'double-plus' loadings show that variate I(1) is most similar to the Pintner General Abilities and Metropolitan Word Meaning tests, and that variate I(2), which correlates .56 with I(1), is most similar to the Stanford Achievement Test in Vocabulary. Moreover, variate II(1) is most similar to the Murphy-Durrell Letter Names test; variate II(2), which correlates .55 with it, is most similar to the Stanford Achievement Tests in Word Reading (grade 1), Paragraph Meaning (grade 1), Spelling (grades 1 and 2), and Word Study Skills (grade 1). This, along with the pattern of the rotated component loadings, suggests that a meaningful labelling of the first dimension which contributes to the between-set relationship might be 'vocabulary', and an appropriate name for the second dimension might be 'spelling'.

The results of the normalized rotation are presented in

Tables A-VI and A-VII and in Figure A-3, and show a pattern very similar to that produced by the raw rotation. The correlations between I(1) and I(2) and between II(1) and II(2), however, are closer in magnitude to the original canonical correlations. This phenomenon is common to all the rotations presented in this paper.

4.3 Ability and achievement tests vs. school grades

(See Appendix B)

In a similar study, Lohnes and Marshall (1965) report a canonical analysis relating a series of educational tests to recorded grades in junior high school. These data were also used by Stewart and Love (1968) in demonstrating the redundancy index.

Lohnes and Marshall report only two pairs of canonical variates. Stewart and Love report all eight pairs. This paper reports three pairs, largely because to do so provides a good demonstration that a 'general factor' situation exists in this analysis as well, but is apparent in three dimensions instead of two. Axes from this analysis were rotated substantially within-sets without disturbing the between-set correlation situation as dramatically as in the Lohnes and Gray study of section 4.2. Moreover, it is interesting to work from the double-plus loadings in Tables B-II and B-IV and relate tests from variable set 1 with school subjects from set 2. Curiously, I(2) turns out to be similar to Arithmetic grades, while II(2) is similar to English grades. A more intuitively satisfying result would have interchanged the roles of these variates.

4.4 16 Personality Factors vs. Vocational Preference Inventory

(See Appendix C)

Williams and Williams (1973) report a canonical analysis relating the 16 Personality Factors test (Cattell, Eber, and Tatsuoka, 1970) to the 11 scales of the Vocational Preference Inventory (Holland, 1965). Tables C-I to C-V and Figures C-3 to C-6 show the results of the Williams study, and the results of a raw and a normalized rotation of the canonical variates. Rotation yields a dramatic increase in the number of very small loadings. Also, the first two variates in each set bear an intuitively simple interpretation after rotation. Variates III(1) and III(2) are oriented somewhat differently by the normalized rotation than by the raw rotation. Note, however, that the canonical correlation associated with these variates is substantially smaller than the largest two canonical correlations. We may, in fact, be trying to rotate one too many variate pairs. Either rotation scheme leaves the earlier two pairs of variates open to much the same interpretation.

The plot of the normalized-rotated loadings shows that this rotation method was able to achieve as much of a sense of 'simple structure' as was the raw rotation, while not affecting the system of variate intercorrelations as much.

4.5 Recalled parental behaviour vs. M.M.P.I. scales

(See Appendix D)

Burger, Armentrout, and Rapfogel (1975) report several canonical analyses between the Child's Report of Parental Behaviour Inventory (Schaefer, 1965) and various objective personality measures. The subjects for the part of their study which is reproduced here are 83 males recalling their father's behaviour, and who took the Minnesota Multiphasic Personality Inventory (Hathaway and McKinley, 1951). This particular canonical analysis is the only one performed in the Burger study which yielded more than one statistically significant canonical correlation (using the Bartlett test), and hence for which more than one canonical variate pair was reported. The matrices from this study are in Tables D-I to D-V and the associated plots are in Figures D-3 to D-6.

Rotation had much less effect on this data set than on others considered in this chapter. The best, and simplest, reason for this lack of effect is that good 'simple structure' could not be found in the data, and a scan of the plots of the loading matrices bears this assertion out somewhat. One noticeable result, however, is the 'cleaning up' of variate II(1). The large loadings on the unrotated II(1), however, seem to have been shifted to I(1) and III(1), which both begin to have the appearance of 'general factors.' Also, the M.M.P.I. scales seem to be somewhat more amenable to rotation than the recalled parental behaviour scales.

4.6 <u>Crime rates from victimization studies vs. crime rates from police sources</u> (See Appendix E)

Scott (1977) reports a canonical analysis relating a set of seven crime rates obtained from National Crime Panel victimization surveys in 26 American cities with seven crime rates derived from those reported by the United States Federal Bureau of Investigation in its Uniform Crime Reports series. The results of this analysis are reproduced in Tables E-I to E-V and Figures E-3 to E-6.

Rotation lends considerable interpretability to the results of this study. Variates I(1) and I(2) define a pair of correlated dimensions distinctly related to Robbery and Auto Theft, apparently two 'validly' measured crime categories. II(1) and II(2) demonstrate the peculiar result that the two methods of reporting crime lead to <u>negatively</u> correlated results when considering the incidence of violent crimes such as murder, rape, and assault. III(1) and III(2) are positioned a bit differently by normalized rotation than by raw rotation, but, as before, this may only be an indication that m, the number of variate pairs, is too large. Notably, the first two pairs of variates are positioned virtually identically regardless of the rotation method used.

5. DISCUSSION

In most cases considered in chapter 4, rotation of canonical variates leads to a more intuitively sensible interpretation of the results of a canonical analysis. Notably, normalized rotation leads to an acceptable degree of interpretability while permitting the between-set correlation structure to remain closer to that from the unrotated solution. This latter result indicates a certain economy in normalized rotation, in that it maintains the optimizing properties of the unrotated solution to a greater degree than does raw rotation.

Common factors in the sense provided by Rao (1965) are usually not estimated in canonical analysis. Carroll (1968), however, has described a situation, similar to that described by Rao, in which a single set of orthogonal variates is derived in canonical analysis. Carroll's approach permits easy generalization to canonical analysis of more than two variable sets: Regardless of the number of sets, the between-set relationships are always described with reference to a single set of canonical variates.

In taking seriously the common-factor approach to canonical analysis, one must recognize the introduction of a concrete rationale for considering identical transformations of matrices of structure correlations between the observed variables and the common factos. In such a case, not the matrices A_1 and

A_2 but the matrices P_1 and P_2 of section 2.3 would be rotated; moreover, a rotation scheme such as Hakstian's (1976) would be more appropriate, since there would no longer be any need to consider canonical correlations in performing the rotations. For the reasons discussed in section 3.3, however, simple structure may be all the more difficult to obtain when the two structure matrices are constrained to be rotated identically.

The 'number of factors' problem deserves some further study. The rotational inconsistencies noted in sections 4.4 and 4.6 may be the result of faulty judgment in setting the value of m, or they may be simply an artifact of a rotation scheme which plays variate correlation against a blind search for 'simple structure'. In particular, a type of rotation different from that considered in this paper may lead to a different notion of the proper number of variate pairs to rotate. In any event, this entire problem would provide a challenging, and potentially useful, avenue for further research.

Appendix A

Data from Lohnes and Gray (1972)

Ability tests vs. achievement tests

Table A-I

Data from Lohnes and Gray (1972)

Unrotated loadings from component analysis of ability and achievement tests

Observed variables	Component loadings	
Set one (ability tests)	I	II
1. Pintner-Cunningham General Abilities	.73 (++)	-.38 (-)
2. Murphy-Durrell Phonemes	.71 (++)	-.19
3. Murphy-Durrell Letter Names	.72 (++)	-.10
4. Murphy-Durrell Learning Rate	.55 (+)	-.28
5. Thurstone Pattern Copy	.54 (+)	-.33 (-)
6. Thurstone Identical Forms	.41 (+)	-.39 (-)
7. Metropolitan Word Meaning	.57 (+)	-.45 (-)
8. Metropolitan Listening	.49 (+)	-.51 (-)
Set two (Stanford Achievement Tests)		
1. Word Reading, Grade 1	.85 (++)	.21
2. Word Reading, Grade 2	.84 (++)	.22
3. Paragraph Meaning, Grade 1	.86 (++)	.20
4. Paragraph Meaning, Grade 2	.87 (++)	.17
5. Vocabulary, Grade 1	.76 (++)	-.18
6. Spelling, Grade 1	.75 (++)	.26
7. Spelling, Grade 2	.79 (++)	.36 (+)
8. Word Study Skills, Grade 1	.84 (++)	.17
9. Word Study Skills, Grade 2	.81 (++)	.18
10. Language, Grade 2	.79 (++)	.16
11. Arithmetic Computation, Grade 2	.64 (+)	.08

Table A-II

Data from Lohnes and Gray (1972)

Varimax rotated loadings from component analysis of ability and achievement tests

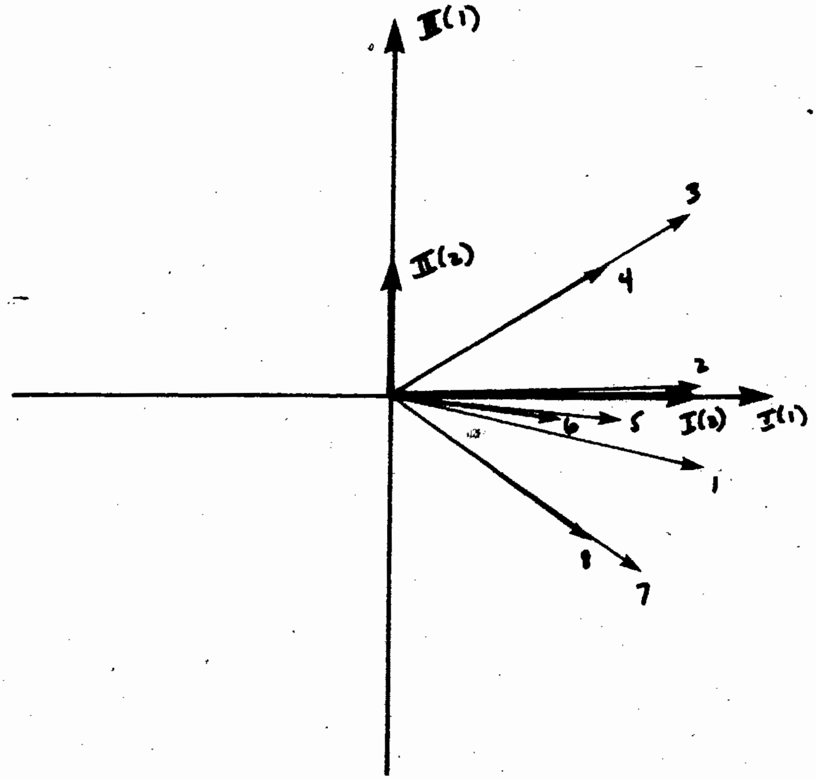
Observed variables	Component loadings			
Set one (ability tests)	I		II	
1. Pintner-Cunningham General Abilities	.37	(+)	.74	(++)
2. Murphy-Durrell Phonemes	.51	(+)	.54	(+)
3. Murphy-Durrell Letter Names	.55	(+)	.47	(+)
4. Murphy-Durrell Learning Rate	.34	(+)	.53	(+)
5. Thurstone Pattern Copy	.30	(+)	.54	(+)
6. Thurstone Identical Forms	.11		.53	(+)
7. Metropolitan Word Meaning	.22		.60	(+)
8. Metropolitan Listening	.11		.71	(++)
Set two (Stanford Achievement Tests)				
1. Word Reading, Grade 1	.82	(++)	.29	
2. Word Reading, Grade 2	.82	(++)	.27	
3. Paragraph Meaning, Grade 1	.85	(++)	.28	
4. Paragraph Meaning, Grade 2	.82	(++)	.33	(+)
5. Vocabulary, Grade 1	.55	(+)	.54	(+)
6. Spelling, Grade 1	.76	(++)	.21	
7. Spelling, Grade 2	.86	(++)	.14	
8. Word Study Skills, Grade 1	.82	(++)	.29	
9. Word Study Skills, Grade 2	.78	(++)	.31	(+)
10. Language, Grade 2	.75	(++)	.31	(+)
11. Arithmetic Computation, Grade 2	.57	(+)	.30	(+)

Table A-III

Data from Lohnes and Gray (1972)

Unrotated loadings from canonical analysis of ability and achievement tests

Observed variables	Canonical loadings			
Set one (ability tests)	I(1)		II(1)	
1. Pintner-Cunningham General Abilities	.84	(++)	-.19	
2. Murphy-Durrell Phonemes	.81	(++)	.02	
3. Murphy-Durrell Letter Names	.79	(++)	.48	(+)
4. Murphy-Durrell Learning Rate	.57	(+)	.34	(+)
5. Thurstone Pattern Copy	.60	(+)	-.06	
6. Thurstone Identical Forms	.44	(+)	-.06	
7. Metropolitan Word Meaning	.67	(+)	-.44	(-)
8. Metropolitan Listening	.54	(+)	-.39	(-)
Set two (Stanford Achievement Tests)	I(2)		II(2)	
1. Word Reading, Grade 1	.84	(++)	.36	(+)
2. Word Reading, Grade 2	.80	(++)	.11	
3. Paragraph Meaning, Grade 1	.84	(++)	.38	(+)
4. Paragraph Meaning, Grade 2	.85	(++)	.12	
5. Vocabulary, Grade 1	.87	(++)	-.31	(-)
6. Spelling, Grade 1	.71	(++)	.50	(+)
7. Spelling, Grade 2	.69	(++)	.47	(+)
8. Word Study Skills, Grade 1	.84	(++)	.20	
9. Word Study Skills, Grade 2	.79	(++)	.09	
10. Language, Grade 2	.76	(++)	.08	
11. Arithmetic Computation, Grade 2	.64	(+)	.11	
Canonical correlations	.81		.31	



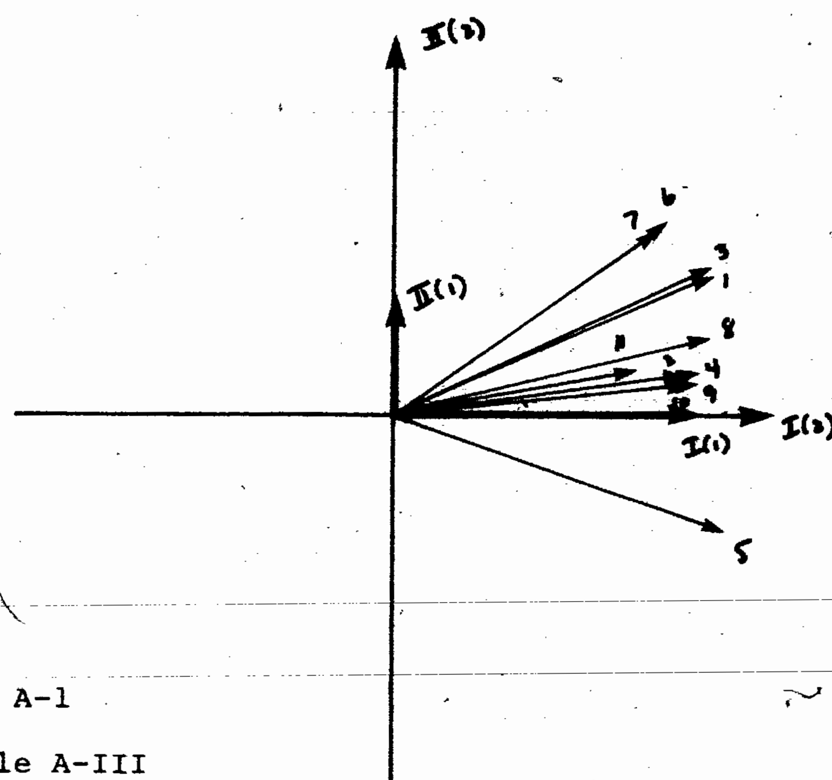


Figure A-1

Plot of Table A-III

Table A-IV

Data from Lohnes and Gray (1972)

Raw rotation of loadings from canonical analysis of ability and achievement tests

Observed variables	Canonical loadings	
Set one (ability tests)	I(1)	II(1)
1. Pintner-Cunningham General Abilities	.72 (++)	.47 (+)
2. Murphy-Durrell Phonemes	.55 (+)	.60 (+)
3. Murphy-Durrell Letter Names	.20	.90 (++)
4. Murphy-Durrell Learning Rate	.15	.65 (+)
5. Thurstone Pattern Copy	.46 (+)	.39 (+)
6. Thurstone Identical Forms	.35 (+)	.28
7. Metropolitan Word Meaning	.78 (++)	.18
8. Metropolitan Listening	.66 (+)	.12
Set two (Stanford Achievement Tests)	I(2)	II(2)
1. Word Reading, Grade 1	.37 (+)	.83 (++)
2. Word Reading, Grade 2	.51 (+)	.62 (+)
3. Paragraph Meaning, Grade 1	.36 (+)	.85 (++)
4. Paragraph Meaning, Grade 2	.54 (+)	.66 (+)
5. Vocabulary, Grade 1	.85 (++)	.36 (+)
6. Spelling, Grade 1	.18	.85 (++)
7. Spelling, Grade 2	.19	.81 (++)
8. Word Study Skills, Grade 1	.48 (+)	.72 (++)
9. Word Study Skills, Grade 2	.52 (+)	.60 (+)
10. Language, Grade 2	.50 (+)	.57 (+)
11. Arithmetic Computation, Grade 2	.40 (+)	.51 (+)

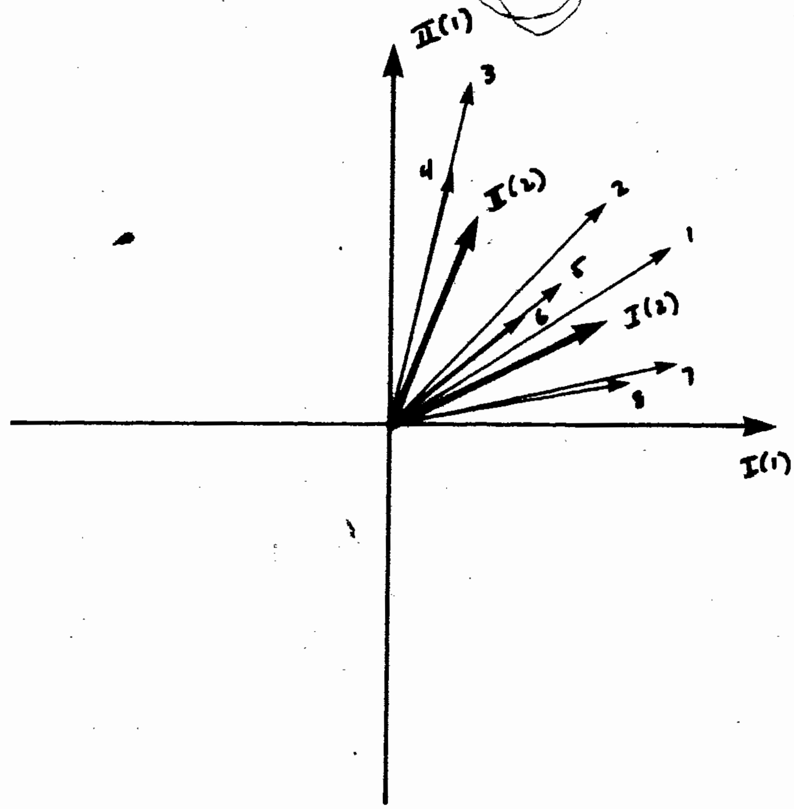
Table A-V

Data from Lohnes and Gray (1972)

Variate intercorrelations from canonical analysis of ability and achievement tests

Raw rotation

		Achievement test variates	
		I (2)	II (2)
Ability test variates	I (1)	.56	.22
	II (1)	.28	.55



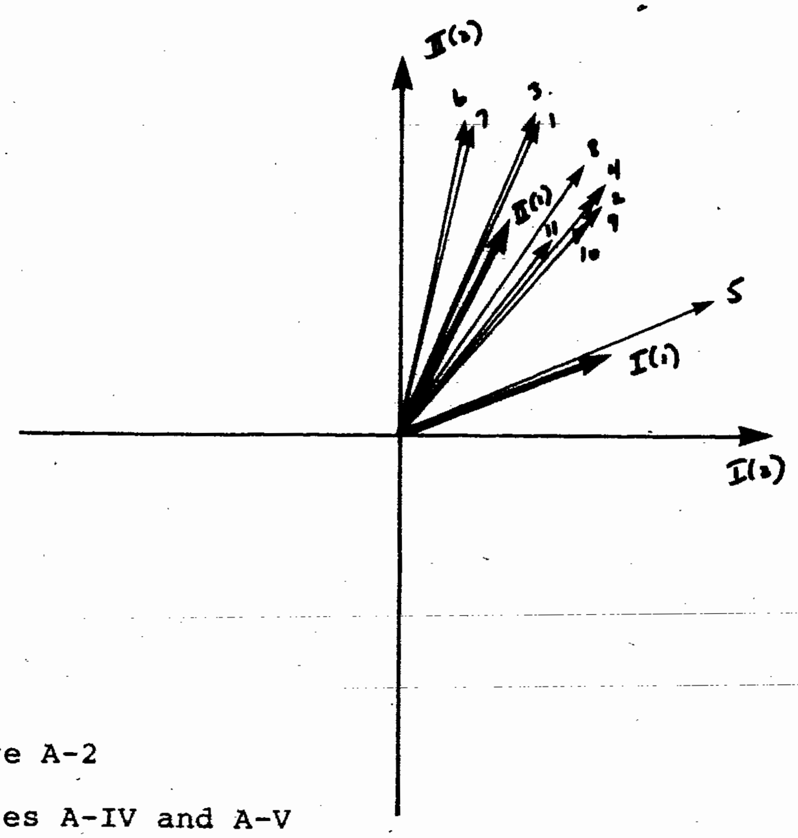


Figure A-2

Plot of Tables A-IV and A-V

Table A-VI

Data from Lohnes and Gray (1972)

Normalized rotation of loadings from canonical analysis of ability and achievement tests

Observed variables	Canonical loadings	
Set one (ability tests)	I(1)	II(1)
1. Pintner-Cunningham General Abilities	.75 (++)	.42 (+)
2. Murphy-Durrell Phonemes	.59 (+)	.56 (+)
3. Murphy-Durrell Letter Names	.27	.89 (++)
4. Murphy-Durrell Learning Rate	.20	.63 (+)
5. Thurstone Pattern Copy	.49 (+)	.36 (+)
6. Thurstone Identical Forms	.37 (+)	.25
7. Metropolitan Word Meaning	.79 (++)	.12
8. Metropolitan Listening	.66 (+)	.07
Set two (Stanford Achievement Tests)	I(2)	II(2)
1. Word Reading, Grade 1	.45 (+)	.79 (++)
2. Word Reading, Grade 2	.57 (+)	.57 (+)
3. Paragraph Meaning, Grade 1	.44 (+)	.81 (++)
4. Paragraph Meaning, Grade 2	.61 (+)	.61 (+)
5. Vocabulary, Grade 1	.88 (++)	.28
6. Spelling, Grade 1	.27	.83 (++)
7. Spelling, Grade 2	.27	.79 (++)
8. Word Study Skills, Grade 1	.55 (+)	.66 (+)
9. Word Study Skills, Grade 2	.58 (+)	.55 (+)
10. Language, Grade 2	.56 (+)	.52 (+)
11. Arithmetic Computation, Grade 2	.45 (+)	.47 (+)

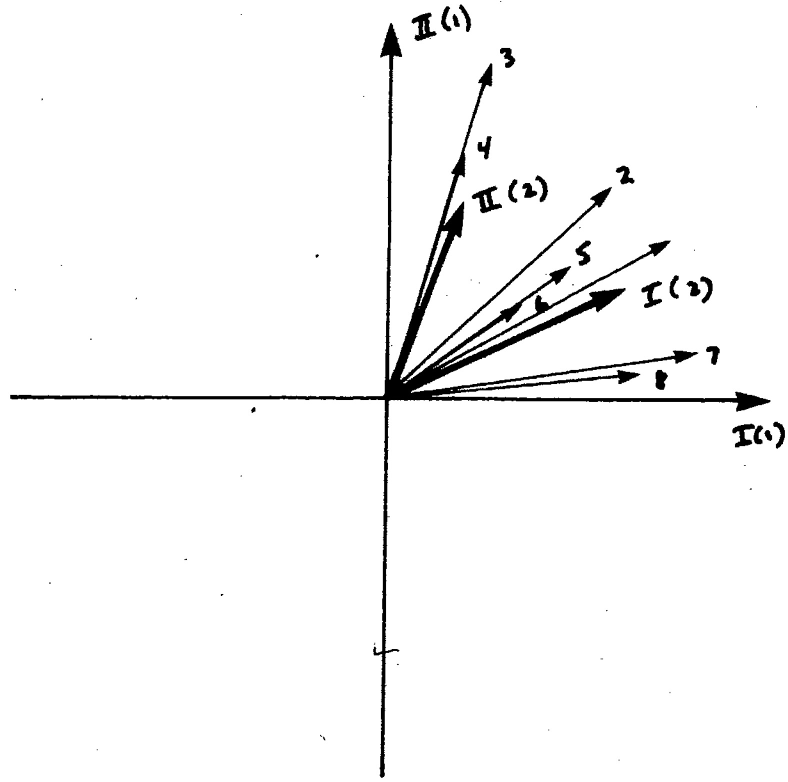
Table A-VII

Data from Lohnes and Gray (1972)

Variate intercorrelations from canonical analysis of ability and achievement tests

Normalized rotation

		Achievement test variates	
		I(2)	II(2)
Ability test variates	I(1)	.61	.20
	II(1)	.29	.51



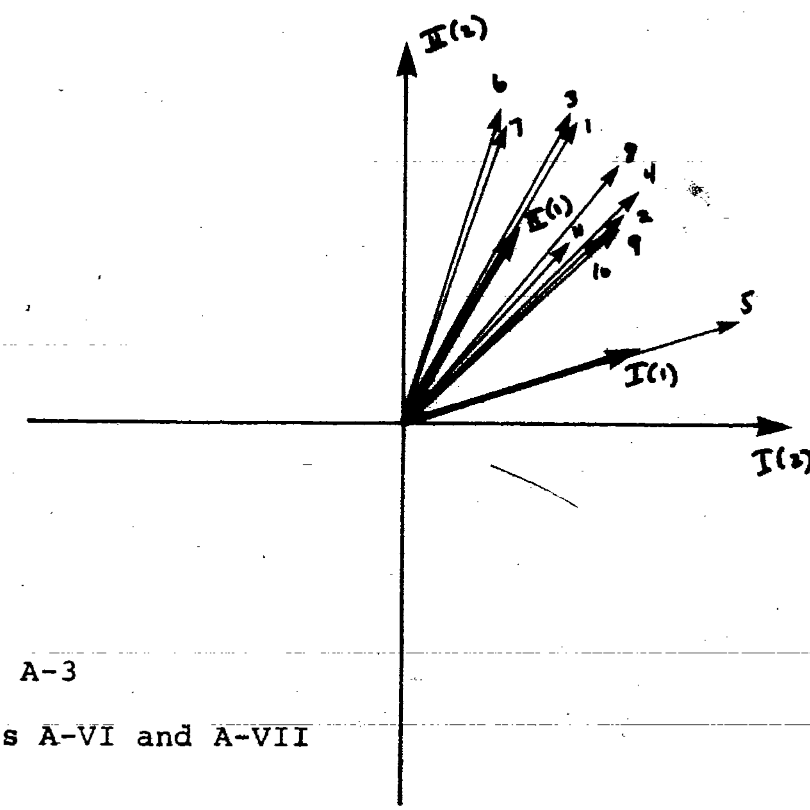


Figure A-3

Plot of Tables A-VI and A-VII

Appendix B

Data from Lohnes and Marshall (1965)

Ability and achievement tests
vs. junior high school grades

Table B-I

Data from Lohnes and Marshall (1965)

Unrotated loadings from canonical analysis
of ability and achievement tests
and junior high school grades

Observed variables	Canonical loadings		
Set one (ability, achievement tests)	I(1)	II(1)	III(1)
1. PGAT Verbal	.79 (++)	-.06	-.08
2. PGAT Reasoning	.83 (++)	.16	.02
3. PGAT Number	.71 (++)	.46 (+)	.01
4. MAT Word Knowledge	.80 (++)	.03	.18
5. MAT Reading	.82 (++)	-.06	.17
6. MAT Spelling	.89 (++)	-.19	-.10
7. MAT Language	.92 (++)	-.12	-.06
8. MAT Study Skills - language	.84 (++)	.07	.09
9. MAT Arithmetic Computation	.90 (++)	.21	-.09
10. MAT Arithmetic Problems	.84 (++)	.35 (+)	.02
11. MAT Social Studies	.75 (++)	-.05	.58 (+)
12. MAT Study Skills - Social studies	.80 (++)	.36 (+)	.14
13. MAT Science	.73 (++)	.19	.22
Set two (junior h. s. grades)	I(2)	II(2)	III(2)
1. 7th Grade English	.85 (++)	.32 (+)	-.07
2. 8th Grade English	.80 (++)	.45 (+)	-.01
3. 7th Grade Arithmetic	.95 (++)	-.14	.01
4. 8th Grade Arithmetic	.88 (++)	-.24	-.01
5. 7th Grade Social Studies	.90 (++)	-.13	.32 (+)
6. 8th Grade Social Studies	.74 (++)	.00	.54 (+)
7. 7th Grade Science	.80 (++)	-.03	.09
8. 8th Grade Science	.73 (++)	.08	.21
Canonical correlations	.90	.66	.50

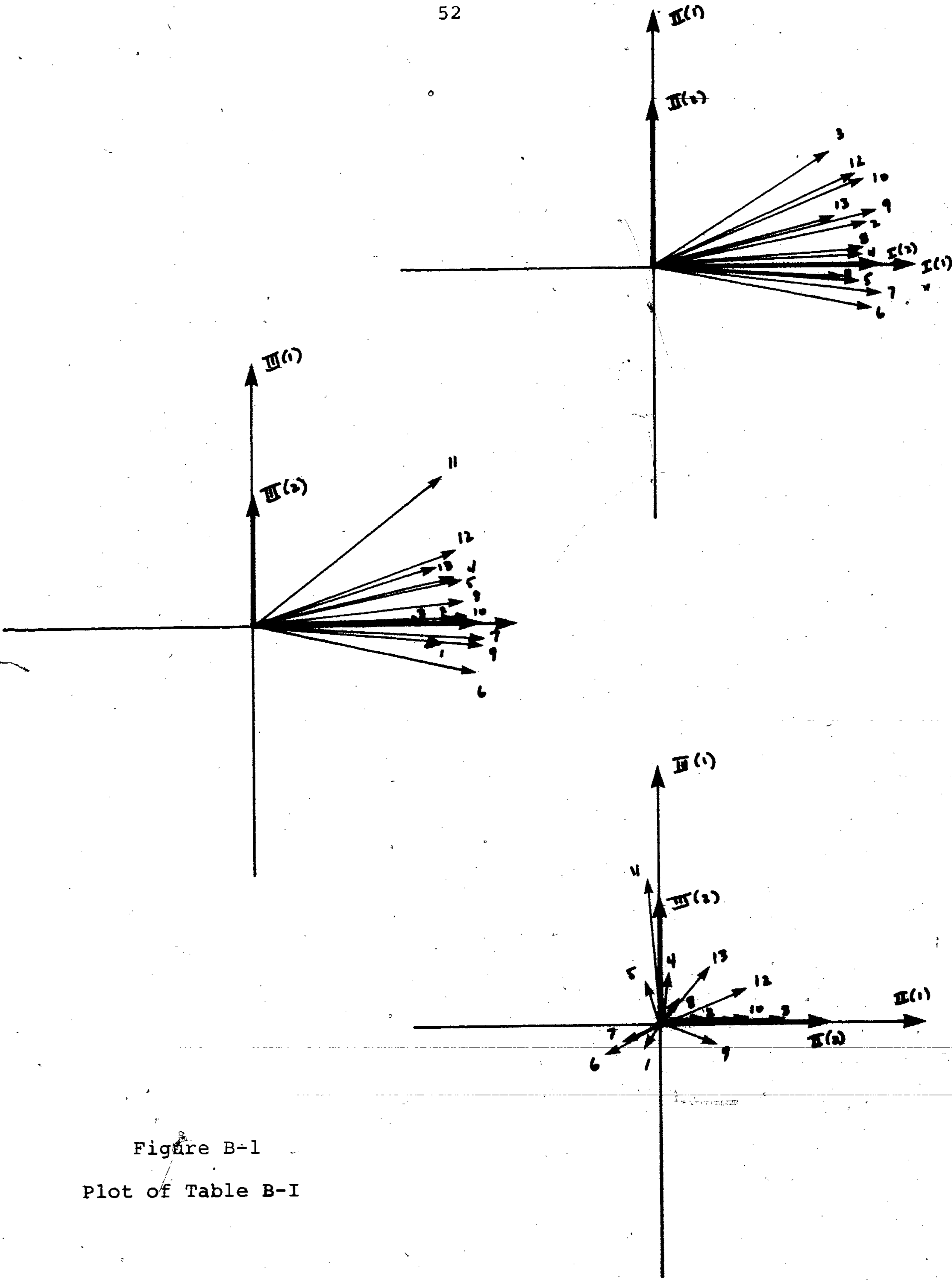
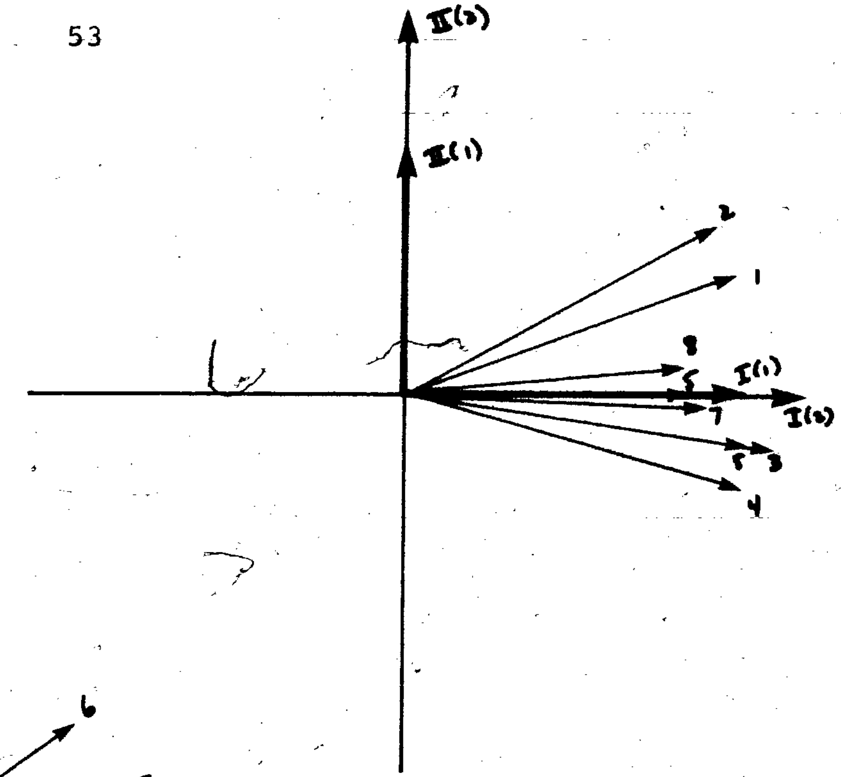
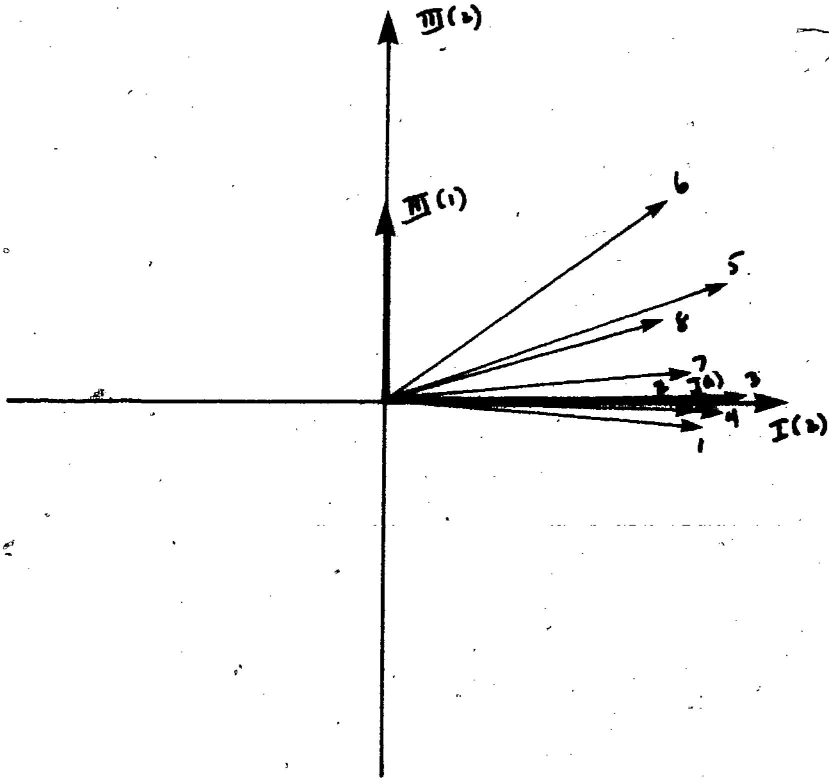


Figure B-1

Plot of Table B-I





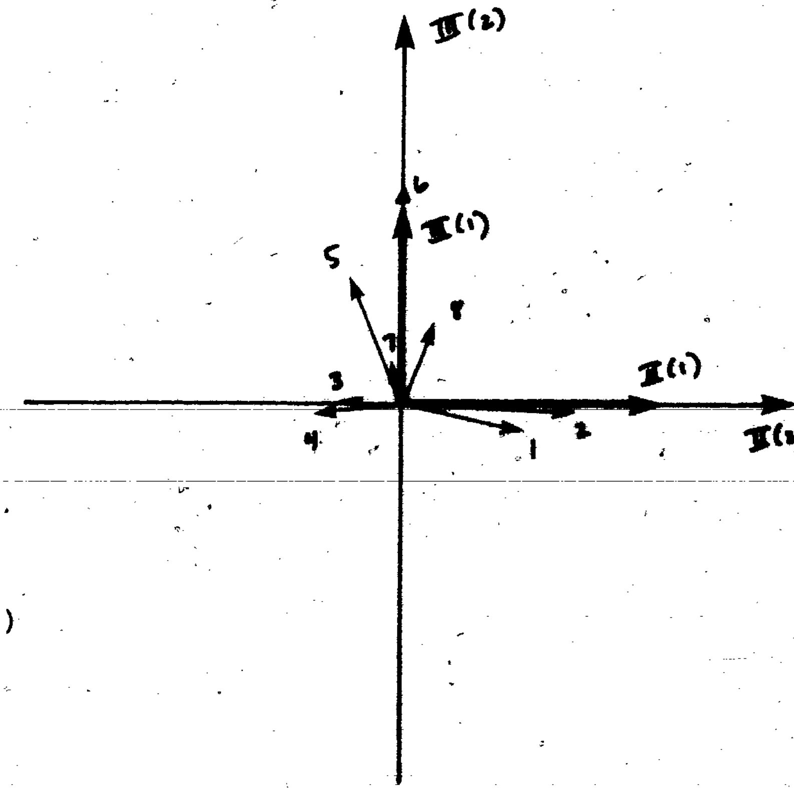


Figure B-1 (cont'd)

Table B-II

Data from Lohnes and Marshall (1965)

Raw rotation of loadings from canonical analysis of ability and achievement tests and junior high school grades

Observed variables	Canonical loadings		
Set one (ability, achievement tests)	I (1)	II (1)	III (1)
1. PGAT Verbal	.67 (+)	.41 (+)	.13
2. PGAT Reasoning	.55 (+)	.60 (+)	.24
3. PGAT Number	.28	.77 (++)	.19
4. MAT Word Knowledge	.57 (+)	.45 (+)	.38 (+)
5. MAT Reading	.63 (+)	.39 (+)	.38 (+)
6. MAT Spelling	.82 (++)	.36 (+)	.15
7. MAT Language	.80 (++)	.43 (+)	.19
8. MAT Study Skills - language	.59 (+)	.51 (+)	.31 (+)
9. MAT Arithmetic Computation	.60 (+)	.70 (++)	.15
10. MAT Arithmetic Problems	.45 (+)	.76 (++)	.23
11. MAT Social Studies	.48 (+)	.31 (+)	.76 (++)
12. MAT Study Skills - Social studies	.38 (+)	.73 (++)	.34 (+)
13. MAT Science	.41 (+)	.54 (+)	.40 (+)
Set two (junior h. s. grades)	I (2)	II (2)	III (2)
1. 7th Grade English	.48 (+)	.76 (++)	.16
2. 8th Grade English	.35 (+)	.82 (++)	.20
3. 7th Grade Arithmetic	.82 (++)	.44 (+)	.24
4. 8th Grade Arithmetic	.83 (++)	.32 (+)	.20
5. 7th Grade Social Studies	.72 (++)	.36 (+)	.52 (+)
6. 8th Grade Social Studies	.49 (+)	.33 (+)	.71 (++)
7. 7th Grade Science	.62 (+)	.42 (+)	.28
8. 8th Grade Science	.48 (+)	.44 (+)	.39 (+)

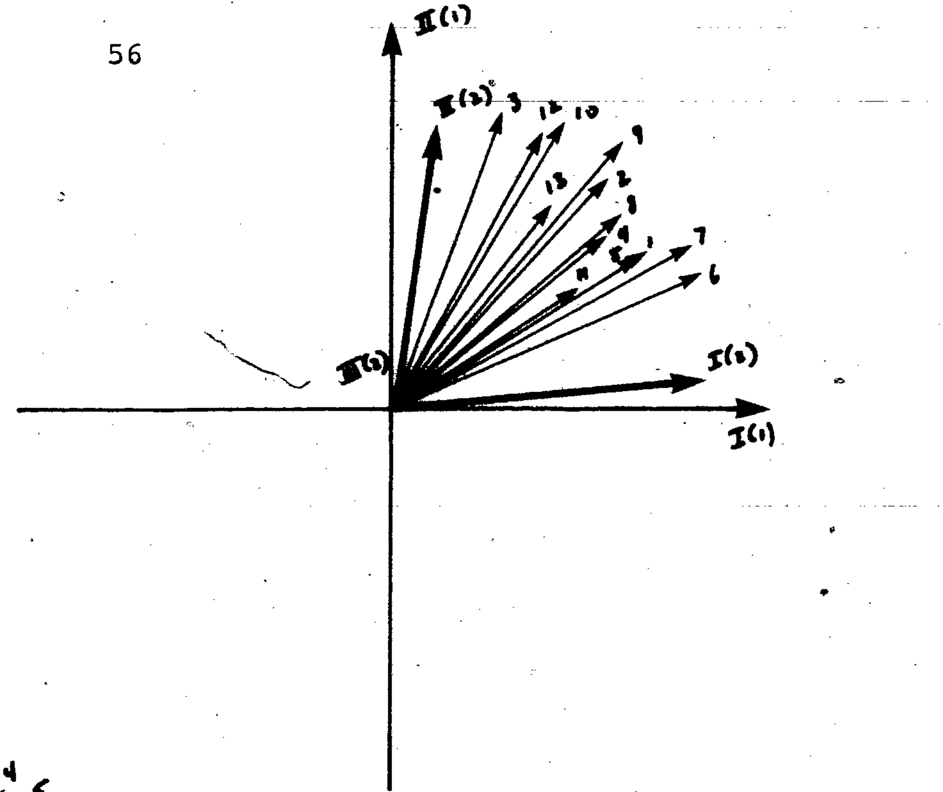
Table B-III

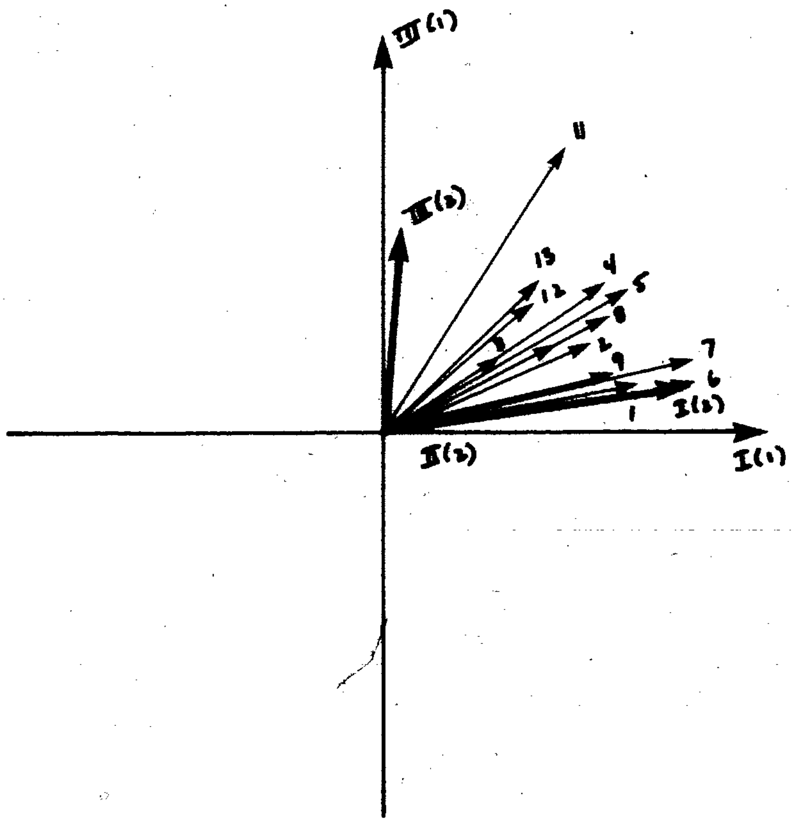
Data from Lohnes and Marshall (1965)

Variate intercorrelations from canonical analysis
of ability and achievement tests
and junior high school grades

Raw rotation

		Junior h. s. grade variates		
		I(2)	II(2)	III(2)
Ability, achievement test variates	I(1)	.80	.12	.05
	II(1)	.08	.74	.08
	III(1)	.11	.04	.53





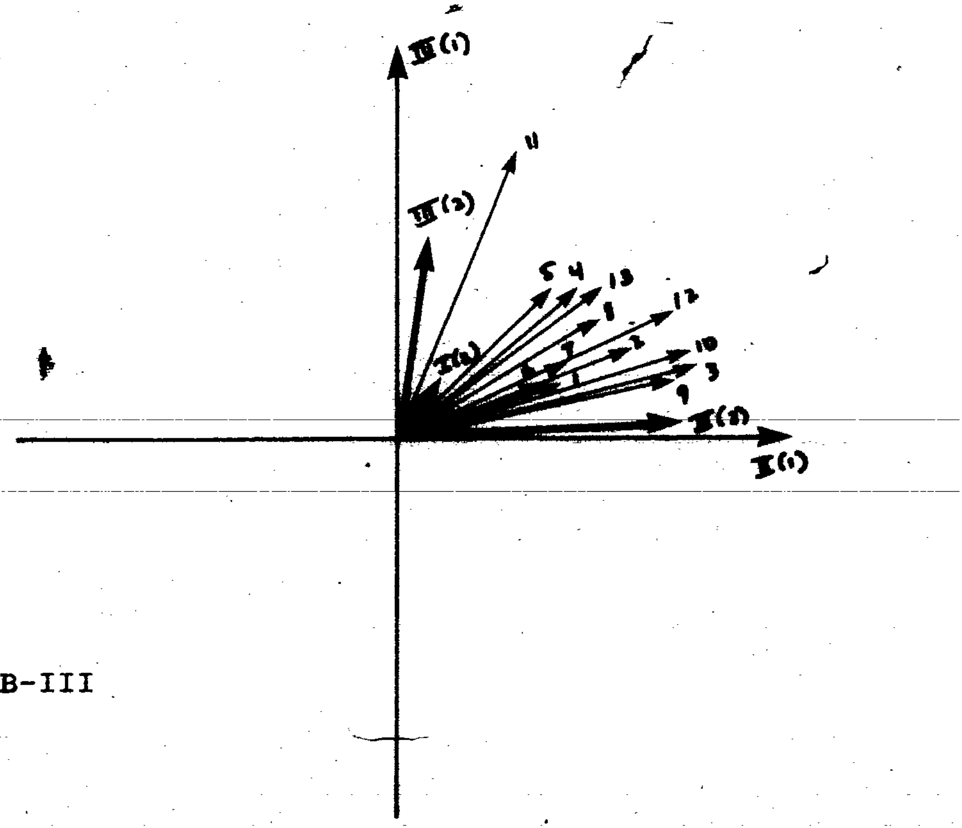
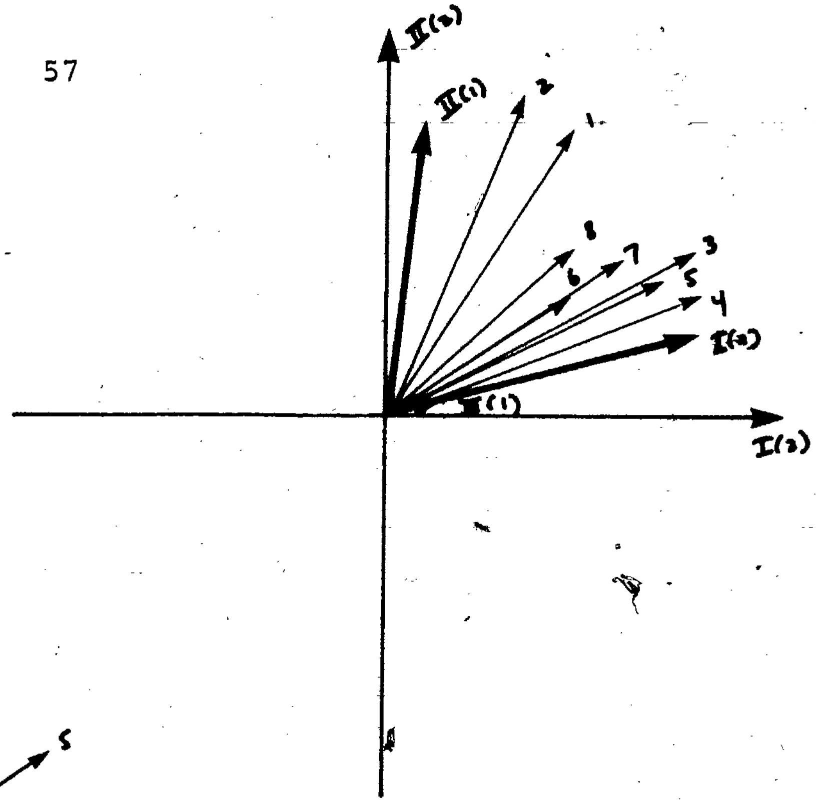
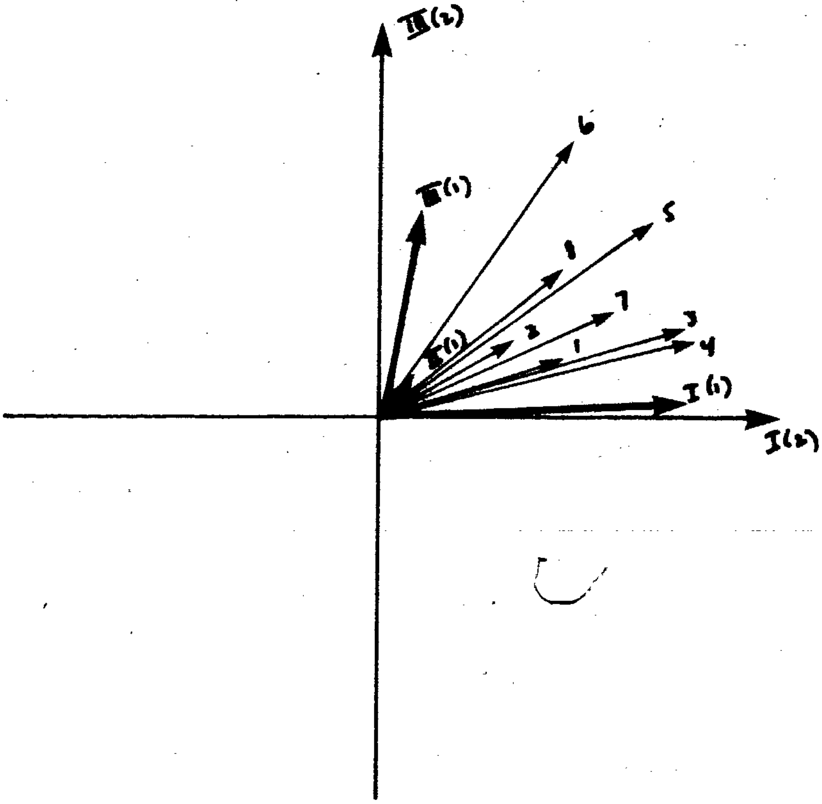


Figure B-2

Plot of Tables B-II and B-III





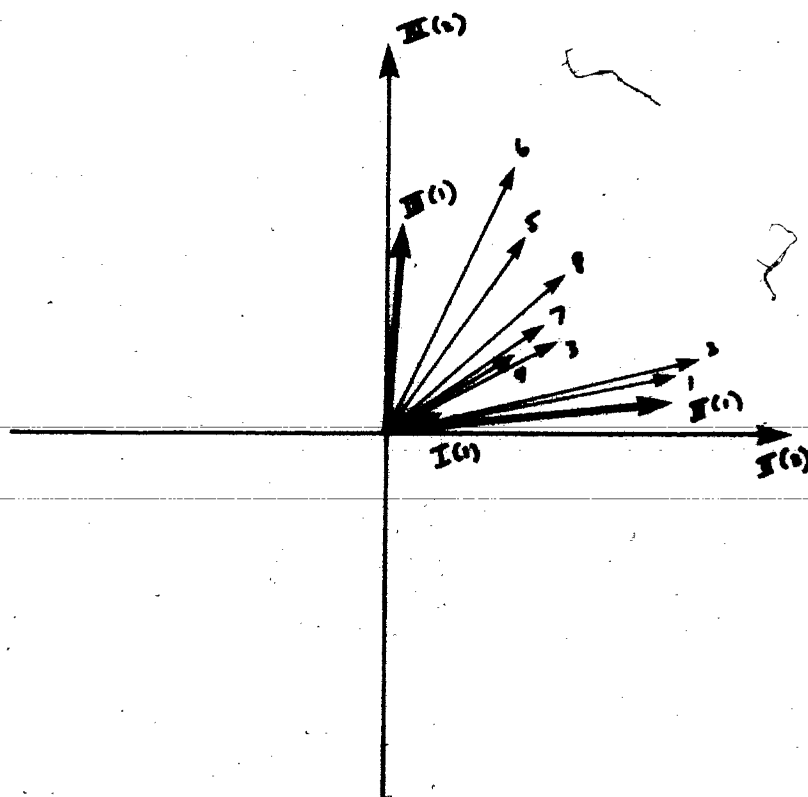


Figure B-2 (cont'd)

Table B-IV

Data from Lohnes and Marshall (1965)

Normalized rotation of loadings from canonical analysis
of ability and achievement tests
and junior high school grades

Observed variables	Canonical loadings		
Set one (ability, achievement tests)	I (1)	II (1)	III (1)
1. PGAT Verbal	.66 (+)	.41 (+)	.13
2. PGAT Reasoning	.54 (+)	.61 (+)	.23
3. PGAT Number	.27	.78 (++)	.17
4. MAT Word Knowledge	.56 (+)	.46 (+)	.38 (+)
5. MAT Reading	.63 (+)	.40 (+)	.38 (+)
6. MAT Spelling	.82 (++)	.37 (+)	.16
7. MAT Language	.79 (++)	.44 (+)	.20
8. MAT Study Skills - language	.59 (+)	.52 (+)	.31 (+)
9. MAT Arithmetic Computation	.60 (+)	.70 (++)	.15
10. MAT Arithmetic Problems	.44 (+)	.77 (++)	.22
11. MAT Social Studies	.47 (+)	.33 (+)	.76 (++)
12. MAT Study Skills - Social studies	.37 (+)	.74 (++)	.33 (+)
13. MAT Science	.40 (+)	.55 (+)	.39 (+)
Set two (junior h. s. grades)	I (2)	II (2)	III (2)
1. 7th Grade English	.46 (+)	.77 (++)	.17
2. 8th Grade English	.33 (+)	.82 (++)	.20
3. 7th Grade Arithmetic	.81 (++)	.45 (+)	.27
4. 8th Grade Arithmetic	.82 (++)	.33 (+)	.22
5. 7th Grade Social Studies	.70 (++)	.38 (+)	.55 (+)
6. 8th Grade Social Studies	.45 (+)	.35 (+)	.72 (++)
7. 7th Grade Science	.61 (+)	.44 (+)	.30
8. 8th Grade Science	.46 (+)	.46 (+)	.40 (+)

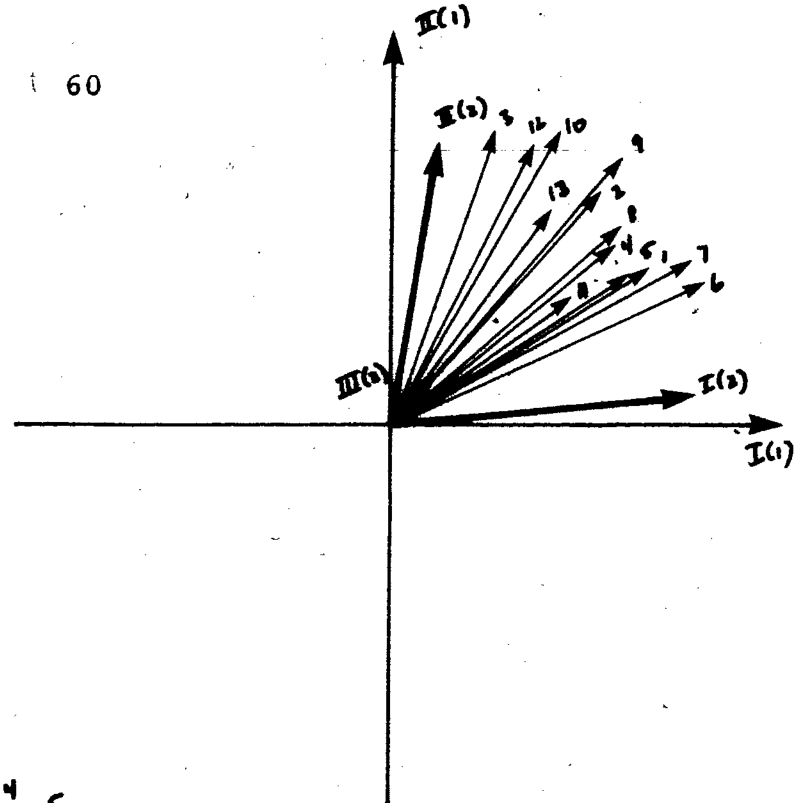
Table B-V

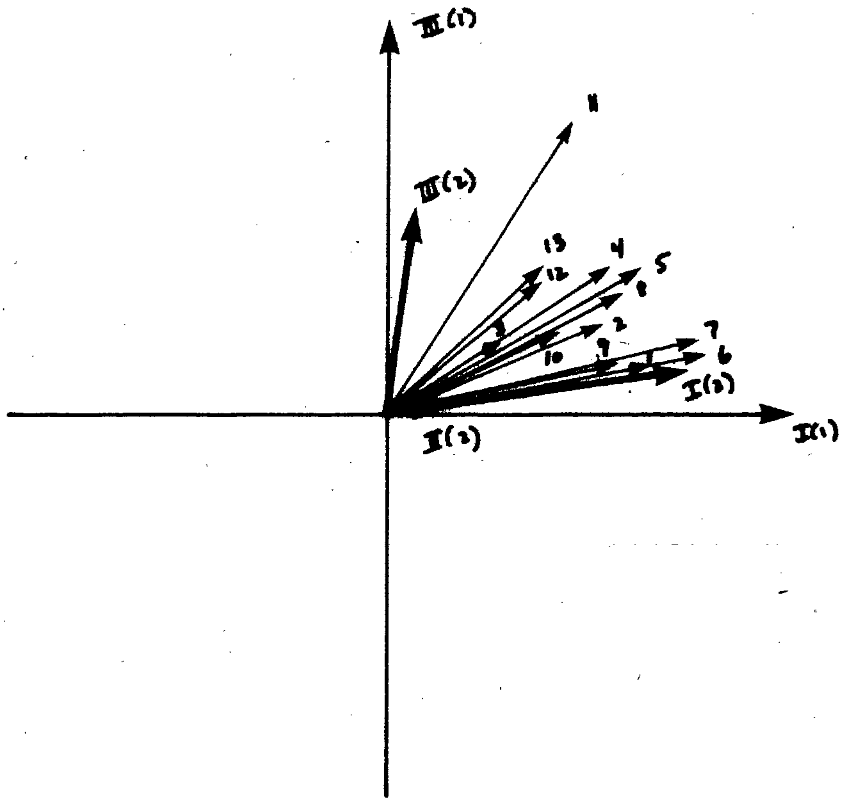
Data from Lohnes and Marshall (1965)

Variate intercorrelations from canonical analysis of ability and achievement tests and junior high school grades

Normalized rotation

		Junior h. s. grade variates		
		I(2)	II(2)	III(2)
Ability, achievement test variates	I(1)	.79	.13	.07
	II(1)	.08	.74	.09
	III(1)	.11	.03	.53





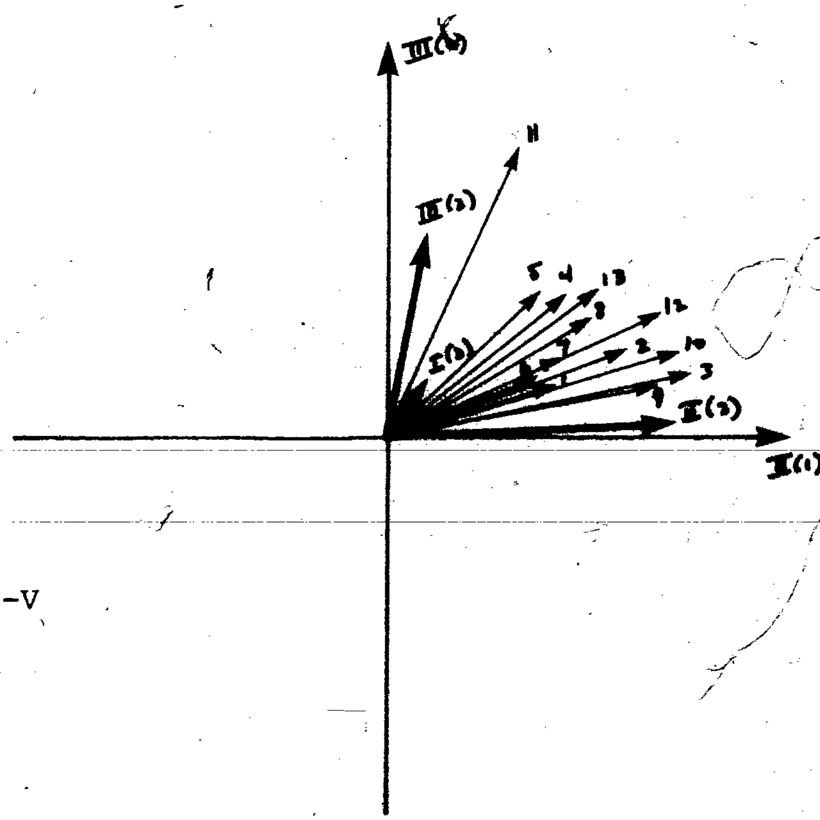
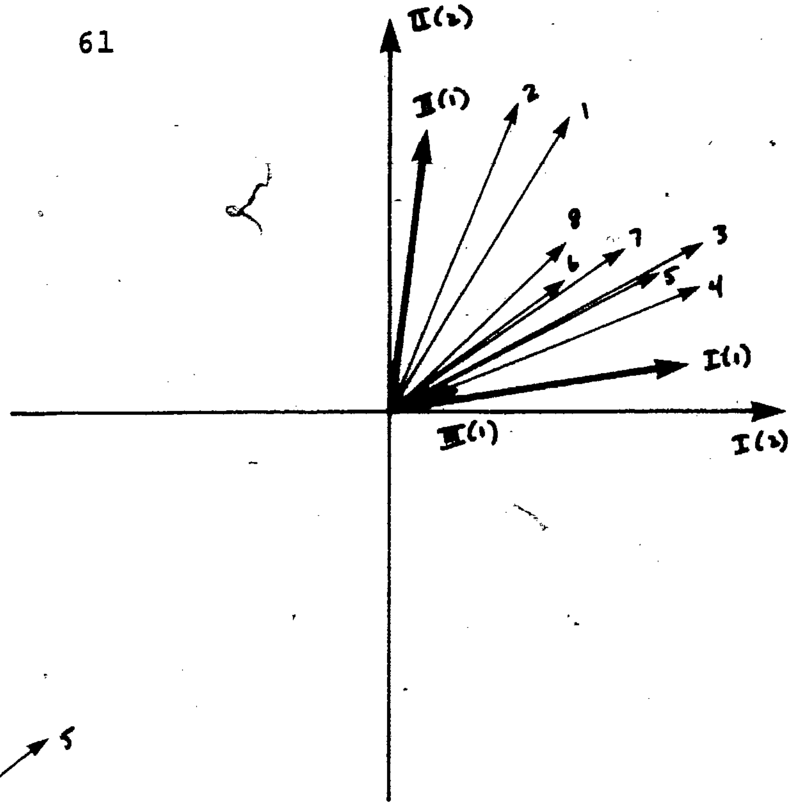
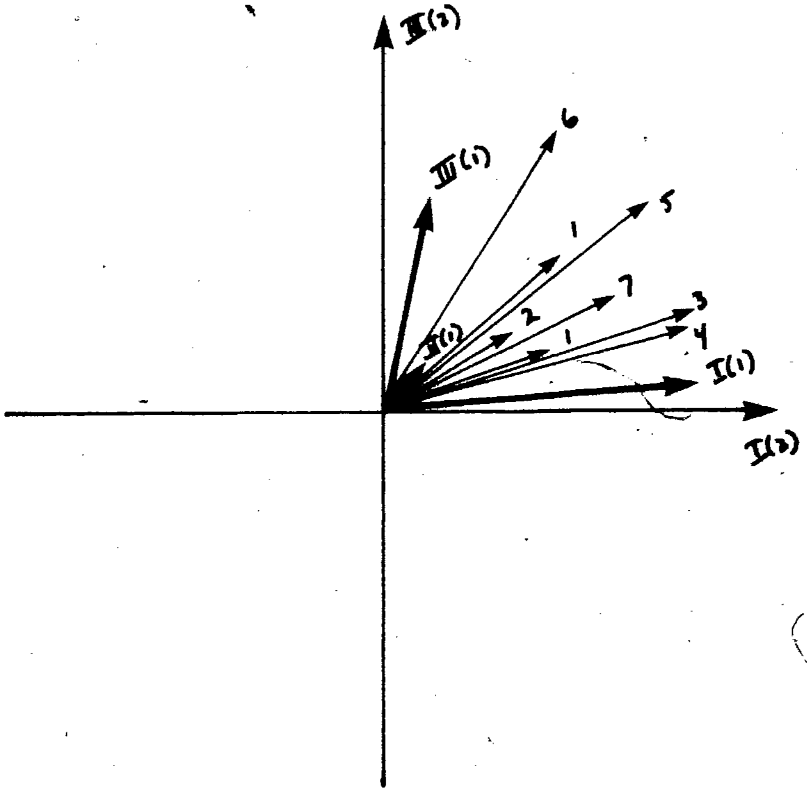


Figure B-3

Plot of Tables B-IV and B-V





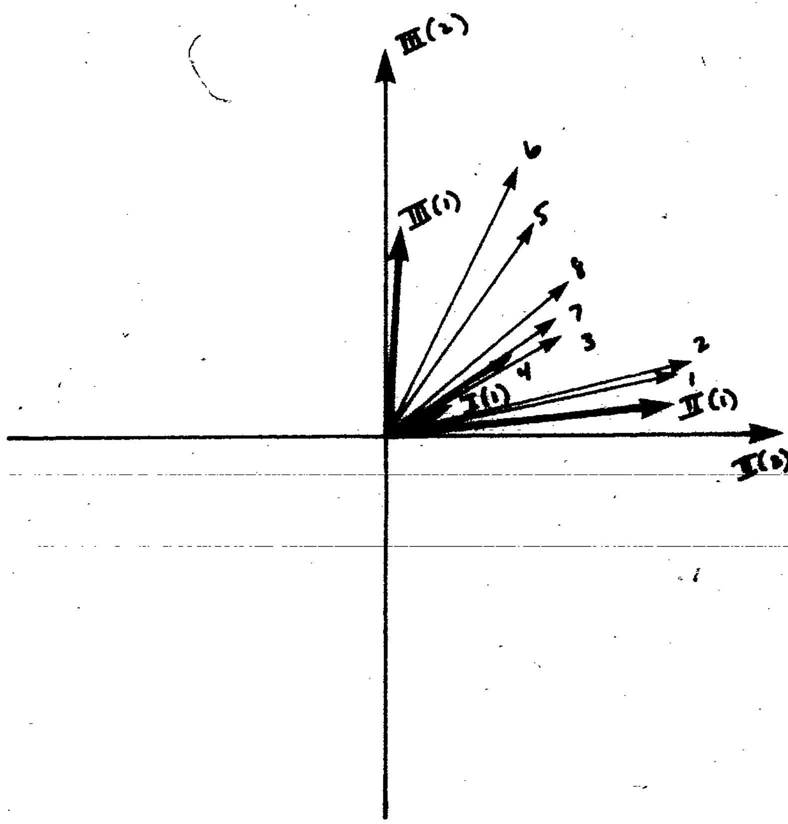


Figure B-3 (cont'd)

Appendix C

Data from Williams and Williams (1973)

16 Personality Factor Questionnaire vs. Vocational Preference Inventory

Table C-I

Data from Williams and Williams (1973)

Unrotated loadings from canonical analysis of the 16 Personality Factor Questionnaire and the Vocational Preference Inventory

Observed variables	Canonical loadings		
Set one (16 P. F.)	I(1)	II(1)	III(1)
1. Cyclothymia (Sociable)	.63 (+)	.58 (+)	.00
2. Intelligence (Bright)	-.02	-.19	.11
3. Emotional Stability (Mature)	.00	.10	.14
4. Dominance (Aggressive)	.07	.21	-.44 (-)
5. Surgency (Enthusiastic)	.28	.53 (+)	-.34 (-)
6. Super-ego Strength (Persistent)	.07	.20	.46 (+)
7. Parmia (Adventurous)	.37 (+)	.34 (+)	-.23
8. Premsia (Effeminate)	.74 (++)	-.43 (-)	-.07
9. Paranoid Tendency (Suspecting)	-.19	-.03	.11
10. Autia (Introverted)	.31 (+)	-.55 (-)	-.27
11. Shrewdness (Sophisticated)	-.15	.08	-.21
12. Guilt Proneness (Insecure)	-.06	-.02	.08
13. Radicalism (Q1)	-.05	-.29	-.18
14. Self-sufficiency (Q2)	-.29	-.62 (-)	.22
15. High Self-sentiment (Q3)	-.13	.04	.04
16. Ergic Tension (Q4)	.07	-.21	.10
Set two (V. P. I.)	I(2)	II(2)	III(2)
1. Realistic	-.33 (-)	-.03	.42 (+)
2. Intellectual (Investigative)	-.29	-.32 (-)	.57 (+)
3. Social	.72 (++)	.40 (+)	.35 (+)
4. Conventional	-.16	.39 (+)	.37 (+)
5. Enterprising	.18	.60 (+)	-.11
6. Artistic	.45 (+)	-.56 (-)	-.21
7. Control	.37 (+)	.10	-.12
8. Male-Female	-.52 (-)	.26	.14
9. Status	.47 (+)	.28	-.22
10. Infrequency	.33 (+)	.20	.06
11. Acquiescence	.07	.04	.28
Canonical correlations	.81	.73	.54

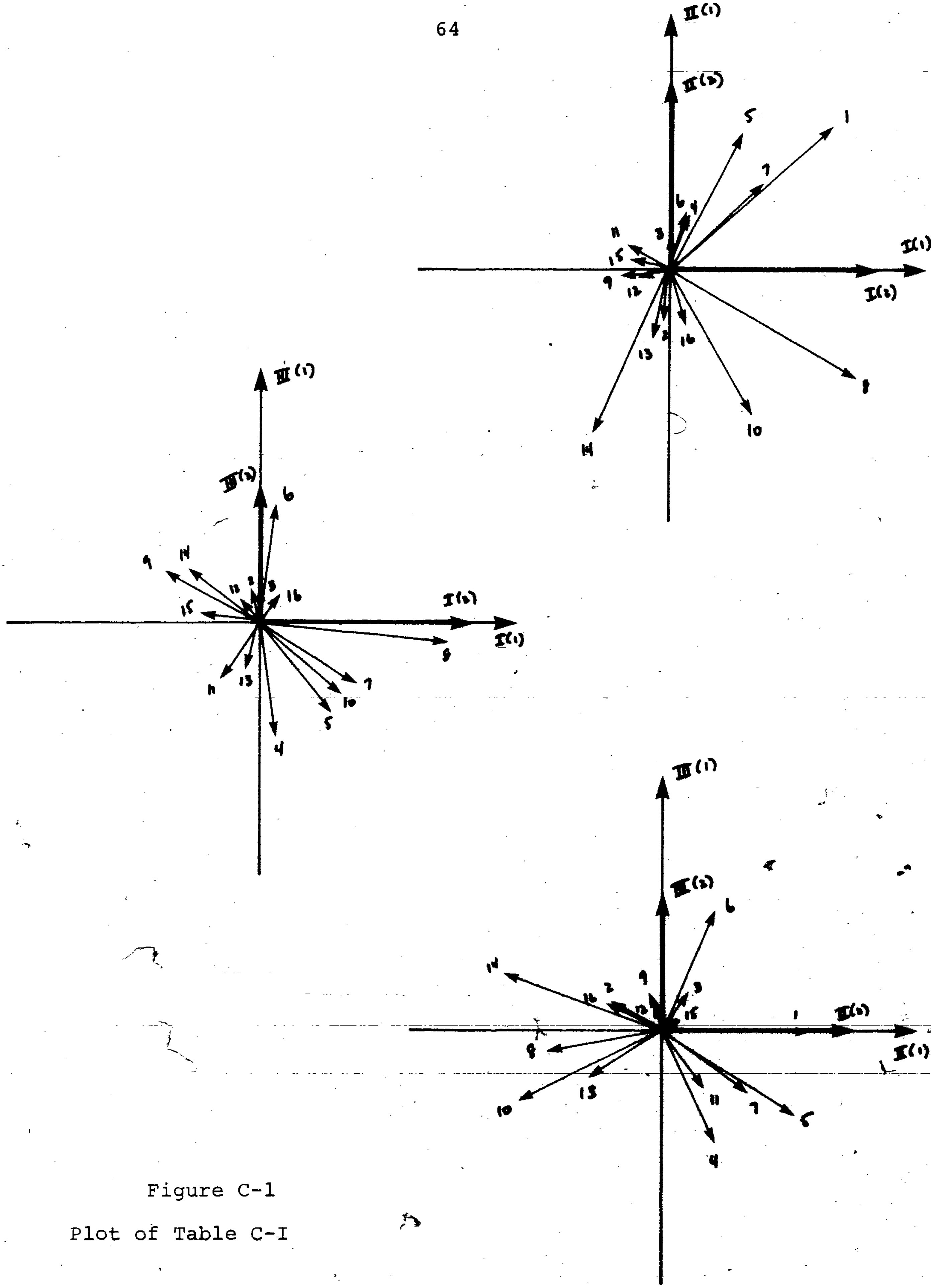


Figure C-1

Plot of Table C-I

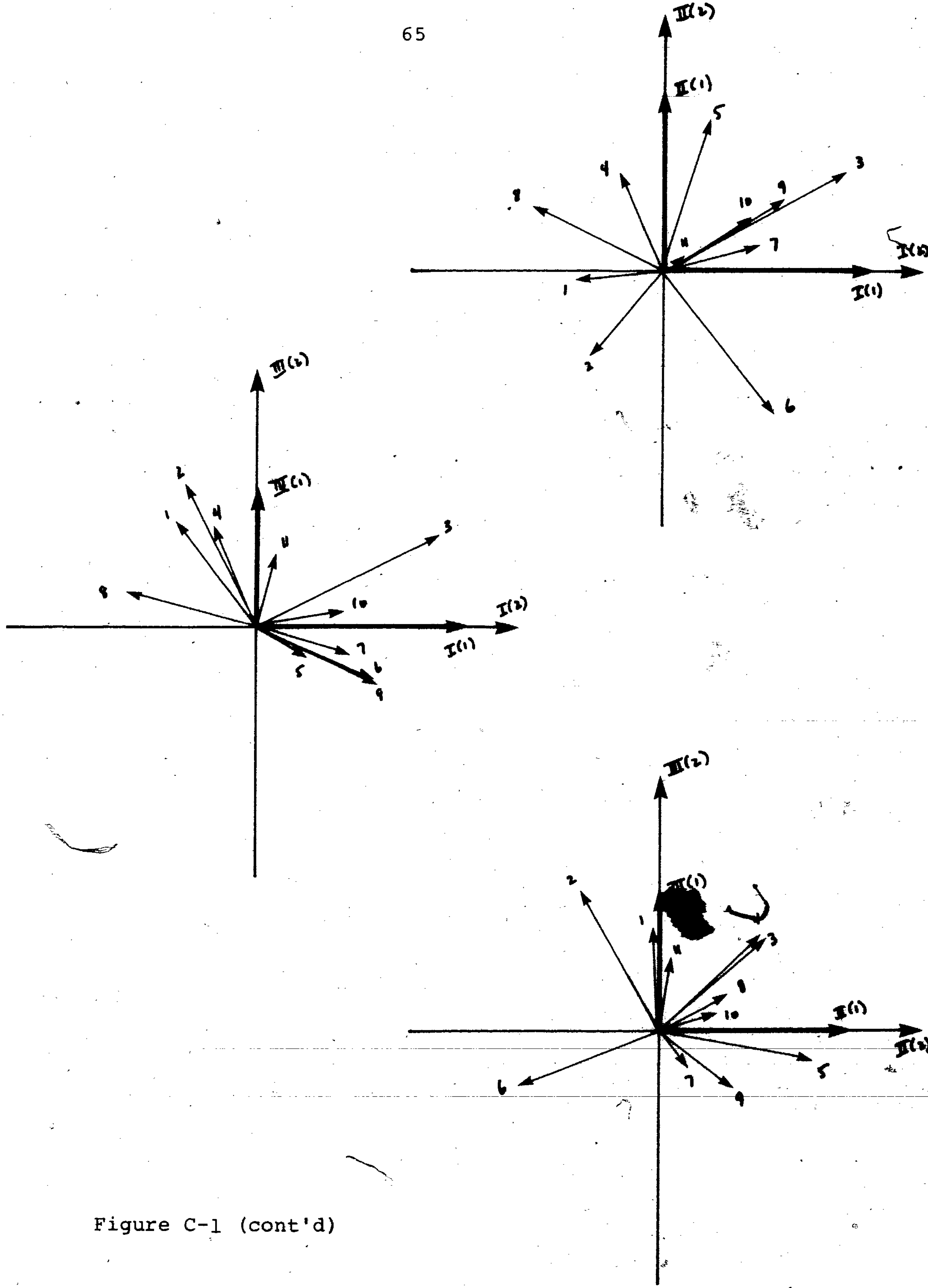


Figure C-1 (cont'd)

Table C-II

Data from Williams and Williams (1973)

Raw rotation of loadings from canonical analysis of the 16 Personality Factor Questionnaire and the Vocational Preference Inventory

Observed variables	Canonical loadings		
Set one (16 P. F.)	I(1)	II(1)	III(1)
1. Cyclothymia (Sociable)	.03	.85 (++)	-.06
2. Intelligence (Bright)	.10	-.14	.14
3. Emotional Stability (Mature)	-.09	.08	.12
4. Dominance (Aggressive)	-.02	.16	-.46 (-)
5. Surgency (Enthusiastic)	-.12	.54 (+)	-.41 (-)
6. Super-ego Strength (Persistent)	-.17	.22	.42 (+)
7. Parmia (Adventurous)	.06	.48 (+)	-.26
8. Premsia (Effeminate)	.83 (++)	.23	.06
9. Paranoid Tendency (Suspecting)	-.13	-.15	.10
10. Autia (Introverted)	.65 (+)	-.18	-.15
11. Shrewdness (Sophisticated)	-.12	-.07	-.23
12. Guilt Proneness (Insecure)	-.04	-.05	.08
13. Radicalism (Q1)	.20	-.25	-.13
14. Self-sufficiency (Q2)	.19	-.62 (-)	.31 (+)
15. High Self-sentiment (Q3)	-.13	-.06	.02
16. Ergic Tension (Q4)	.18	-.09	.14
Set two (V. P. I.)	I(2)	II(2)	III(2)
1. Realistic	-.27	-.16	.43 (+)
2. Intellectual (Investigative)	-.06	-.26	.66 (+)
3. Social	.09	.88 (++)	.14
4. Conventional	-.47 (-)	.20	.23
5. Enterprising	-.31 (-)	.46 (+)	-.32
6. Artistic	.75 (++)	-.04	-.05
7. Control	.19	.31 (+)	-.18
8. Male-Female	-.55 (-)	-.21	.09
9. Status	.14	.46 (+)	-.34 (-)
10. Infrequency	.05	.39 (+)	-.04
11. Acquiescence	-.04	.15	.24

Table C-III

Data from Williams and Williams (1973)

Variate intercorrelations from canonical analysis of the 16 Personality Factor Questionnaire and the Vocational Preference Inventory

Raw rotation

		V. P. I. variates		
		I (2)	II (2)	III (2)
16 P.F. variates	I (1)	.76	-.07	-.01
	II (1)	.02	.73	.05
	III (1)	.01	.11	.58

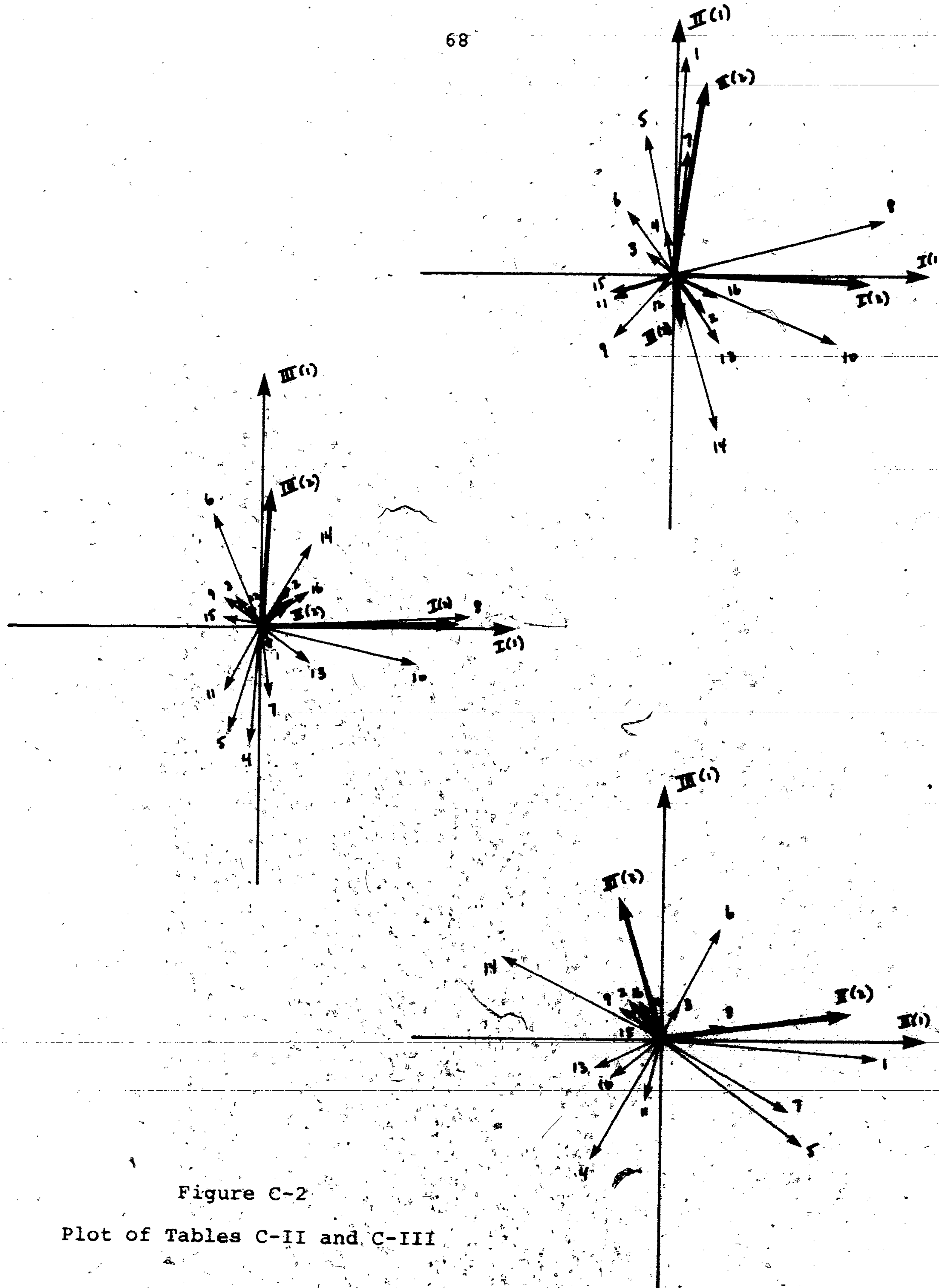
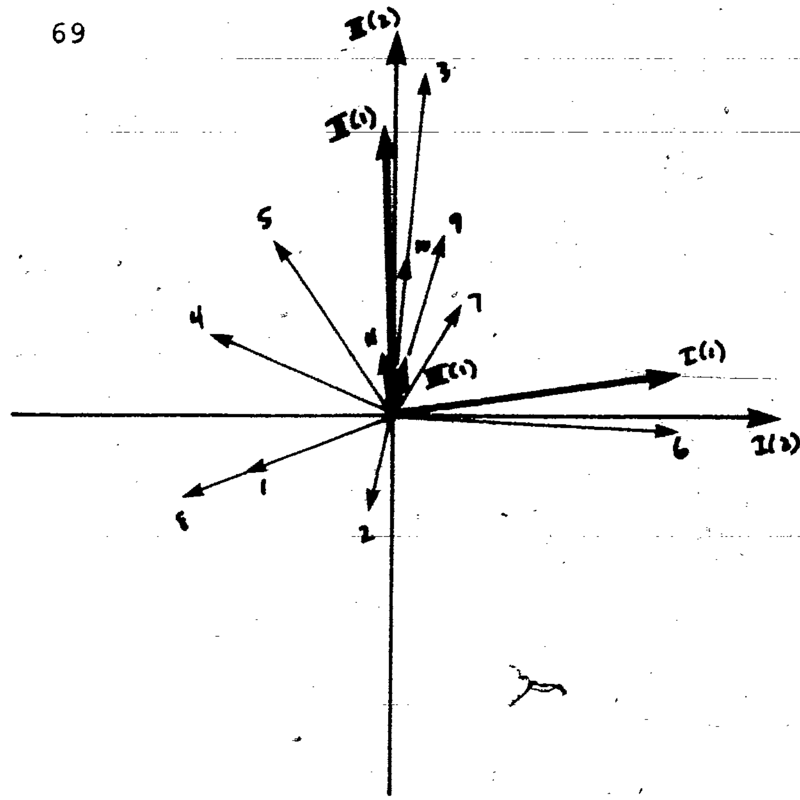
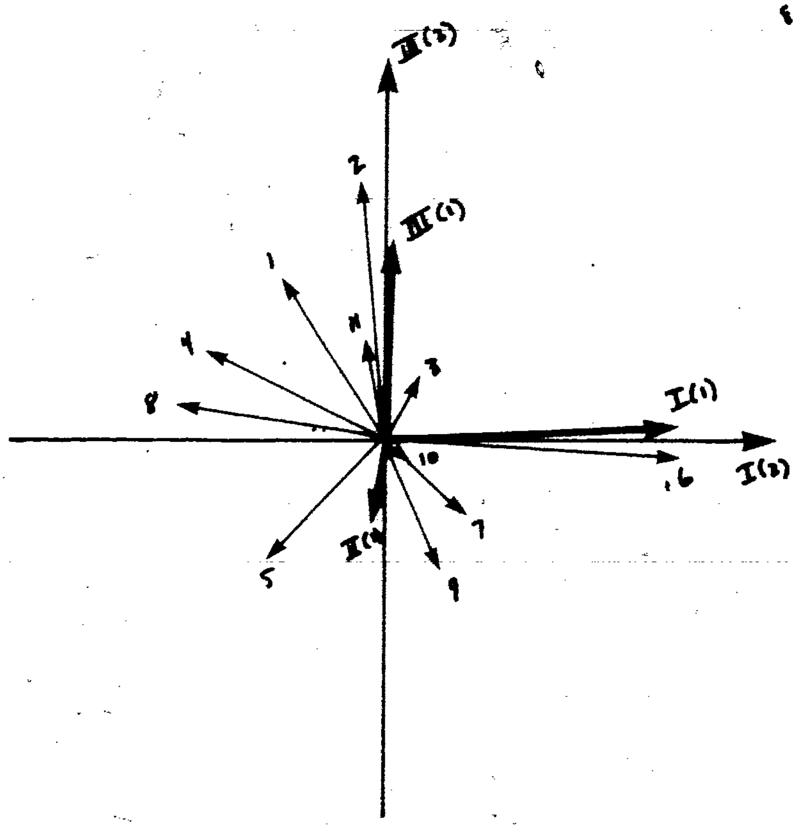


Figure C-2

Plot of Tables C-II and C-III





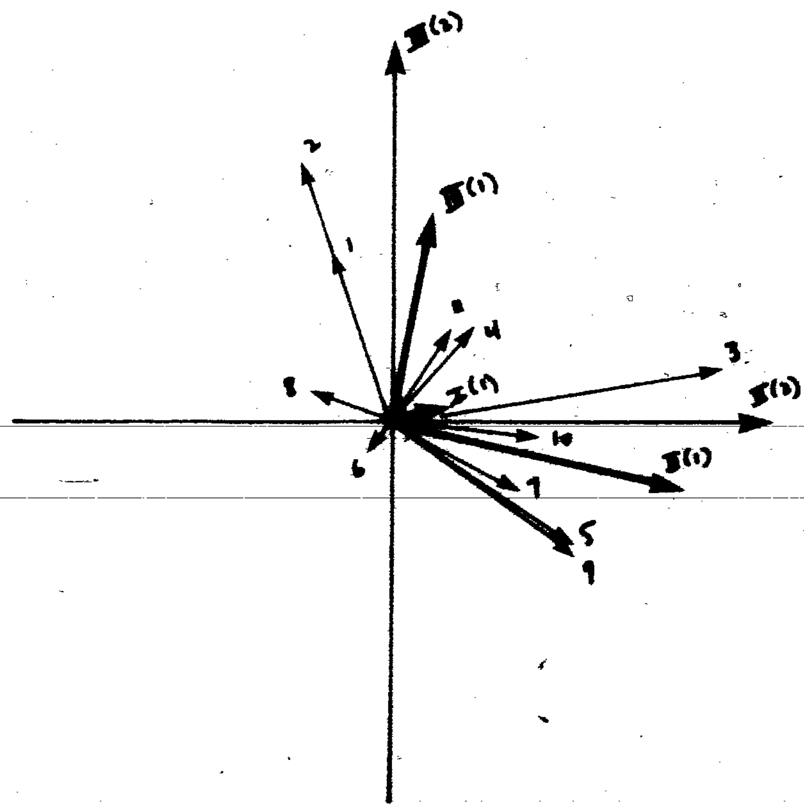


Figure C-2 (cont'd)

Table C-IV

Data from Williams and Williams (1973)

Normalized rotation of loadings from canonical analysis of the 16 Personality Factor Questionnaire and the Vocational Preference Inventory

Observed variables	Canonical loadings		
Set one (16 P. F.)	I (1)	II (1)	III (1)
1. Cyclothymia (Sociable)	.09	.78 (++)	.34 (+)
2. Intelligence (Bright)	.10	-.19	.05
3. Emotional Stability (Mature)	-.08	.02	.15
4. Dominance (Aggressive)	-.03	.36 (+)	-.33 (-)
5. Surgency (Enthusiastic)	-.10	.67 (+)	-.10
6. Super-ego Strength (Persistent)	-.13	.01	.49 (+)
7. Parmia (Adventurous)	.08	.55 (+)	-.01
8. Premsia (Effeminate)	.84 (++)	.14	.09
9. Paranoid Tendency (Suspecting)	-.14	-.17	.03
10. Autia (Introverted)	.62 (+)	-.12	-.26
11. Shrewdness (Sophisticated)	-.14	.05	-.23
12. Guilt Proneness (Insecure)	-.04	-.08	.05
13. Radicalism (Q1)	.17	-.17	-.24
14. Self-sufficiency (Q2)	.16	-.70 (--)	-.03
15. High Self-sentiment (Q3)	-.13	-.06	.00
16. Ergic Tension (Q4)	.18	-.15	.07
Set two (V. P. I.)	I (2)	II (2)	III (2)
1. Realistic	-.28	-.37 (-)	.26
2. Intellectual (Investigative)	-.07	-.60 (+)	.38 (+)
3. Social	.15	.61 (+)	.64 (+)
4. Conventional	-.45 (-)	.04	.33
5. Enterprising	-.28	.57 (+)	.04
6. Artistic	.74 (++)	-.03	-.10
7. Control	.20	.34 (+)	.04
8. Male-Female	-.56 (-)	-.19	-.03
9. Status	.16	.57 (+)	.01
10. Infrequency	.07	.33	.20
11. Acquiescence	-.03	-.03	.29

Table C-V

Data from Williams and Williams (1973)

Variate intercorrelations from canonical analysis
of the 16 Personality Factor Questionnaire
and the Vocational Preference Inventory

Normalized rotation

		V. P. I. variates		
		I(2)	II(2)	III(2)
16 P.F. variates	I(1)	.76	.11	.04
	II(1)	-.02	.75	-.17
	III(1)	.02	.11	.54

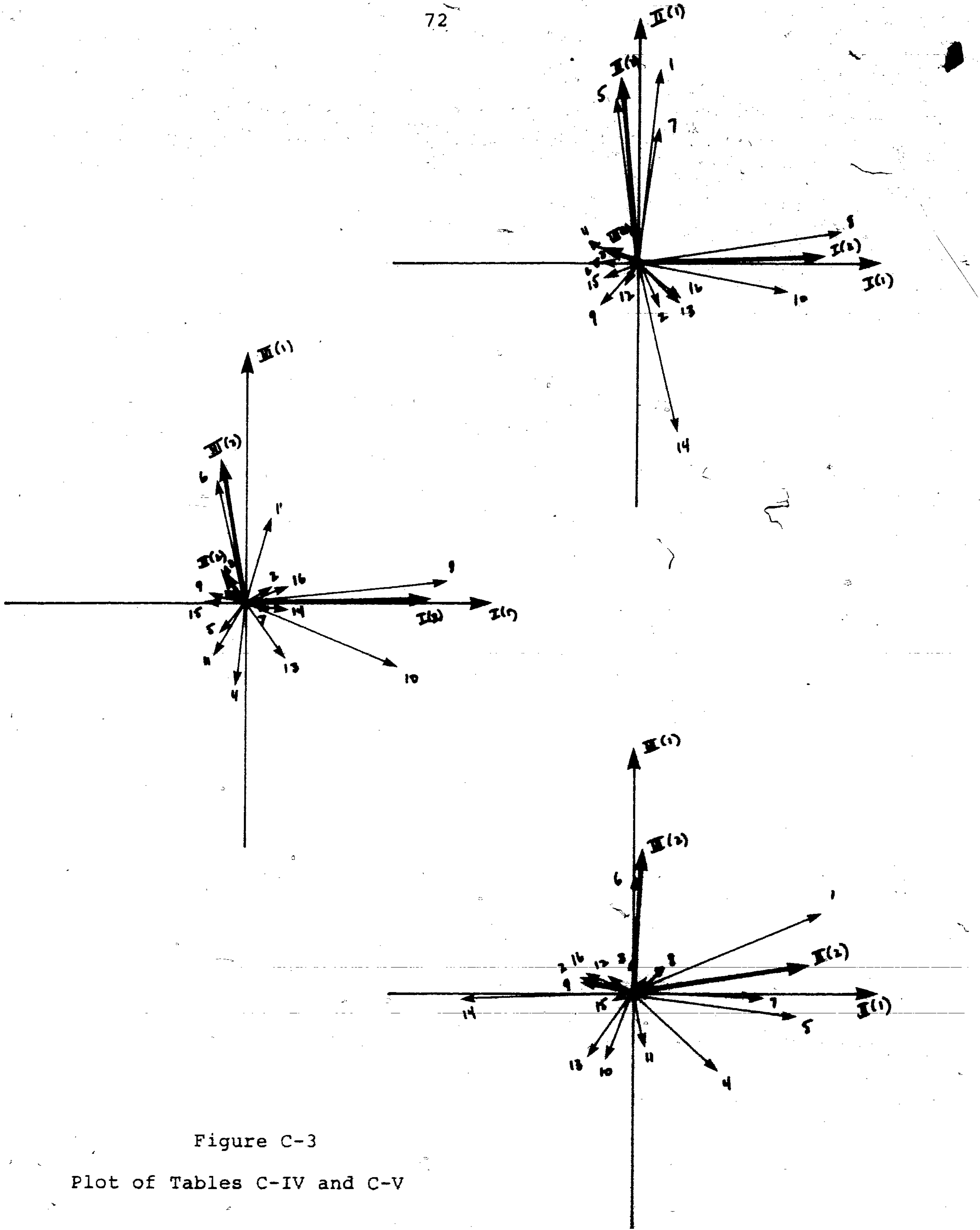


Figure C-3

Plot of Tables C-IV and C-V

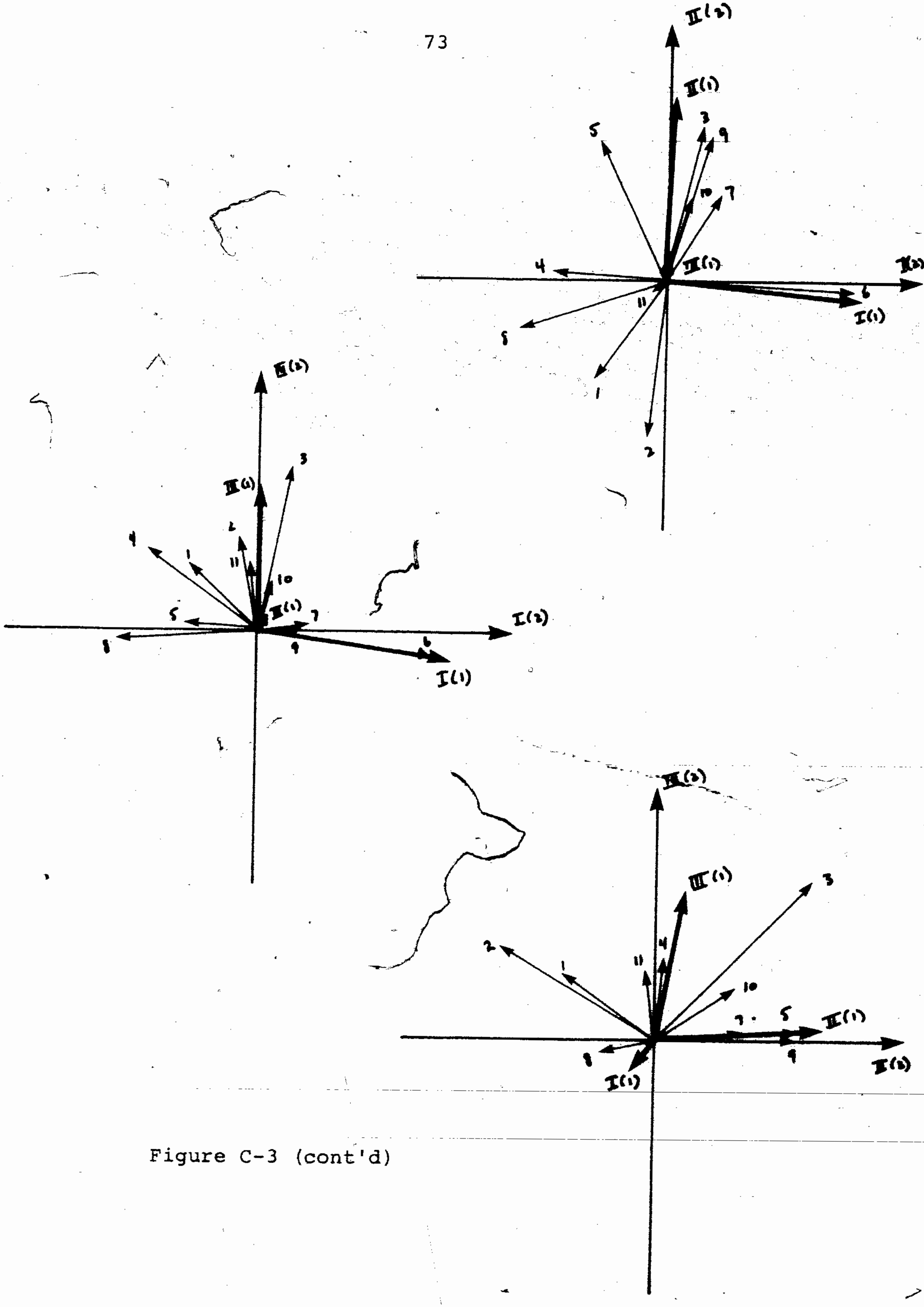


Figure C-3 (cont'd)

Appendix D

Data from Burger, Armentrout, and Rapfogel (1975)

Child's Report of Parental Behavior Inventory vs. Minnesota Multiphasic Personality Inventory

Table D-I

Data from Burger, Armentrout, and Rapfogel (1975)

Unrotated loadings from canonical analysis of Child's Report of Parental Behavior Inventory and Minnesota Multiphasic Personality Inventory

Observed variables	Canonical loadings		
Set one (parental behavior)	I(1)	II(1)	III(1)
1. Acceptance	-.63 (-)	-.32 (-)	-.33 (-)
2. Childcentredness	-.63 (-)	-.11	-.45 (-)
3. Possessiveness	.23	.21	-.07
4. Rejection	.59 (+)	.39 (+)	.40 (+)
5. Control	.10	.26	-.05
6. Enforcement	.43 (+)	.35 (+)	.36 (+)
7. Positive Involvement	-.50 (-)	-.22	-.58 (-)
8. Intrusiveness	.69 (+)	.03	.15
9. Control Through Guilt	.17	.38 (+)	.23
10. Hostile Control	.42 (+)	.56 (+)	.16
11. Inconsistent Discipline	.16	.42 (+)	.27
12. Nonenforcement	.34 (+)	.25	.26
13. Acceptance of Individuation	-.33 (-)	-.37 (-)	-.51 (-)
14. Lax Discipline	.89 (++)	.15	.20
15. Instilling Persistent Anxiety	.03	.40 (+)	.33 (+)
16. Hostile Detachment	.12	.32 (+)	.42 (+)
17. Withdrawl of Relations	.67 (+)	.28	.09
18. Extreme Autonomy	.52 (+)	.22	.23
Set two (M. M. P. I.)	I(2)	II(2)	III(2)
1. L Scale	.89 (++)	-.01	-.02
2. F Scale	.61 (+)	.55 (+)	-.11
3. K Scale	.05	-.42 (-)	.00
4. Hypochondriasis (Hs)	.55 (+)	.46 (+)	-.12
5. Depression (D)	-.26	.08	.58 (+)
6. Hysteria (Hy)	-.33 (-)	.28	.39 (+)
7. Psychopathic Deviate (Pd)	.64 (+)	.65 (+)	.09
8. Masculinity-Femininity (Mf)	-.45 (-)	-.12	.13
9. Paranoia (Pa)	.07	.24	-.24
10. Psychasthenia (Pt)	.07	.40 (+)	-.13
11. Schizophrenia (Sc)	-.20	.62 (+)	-.08
12. Hypomania (Ma)	-.43 (-)	.35 (+)	-.44 (-)
13. Social Introversion-Extroversion	.34 (+)	.20	.36 (+)
Canonical correlations	.98	.81	.75

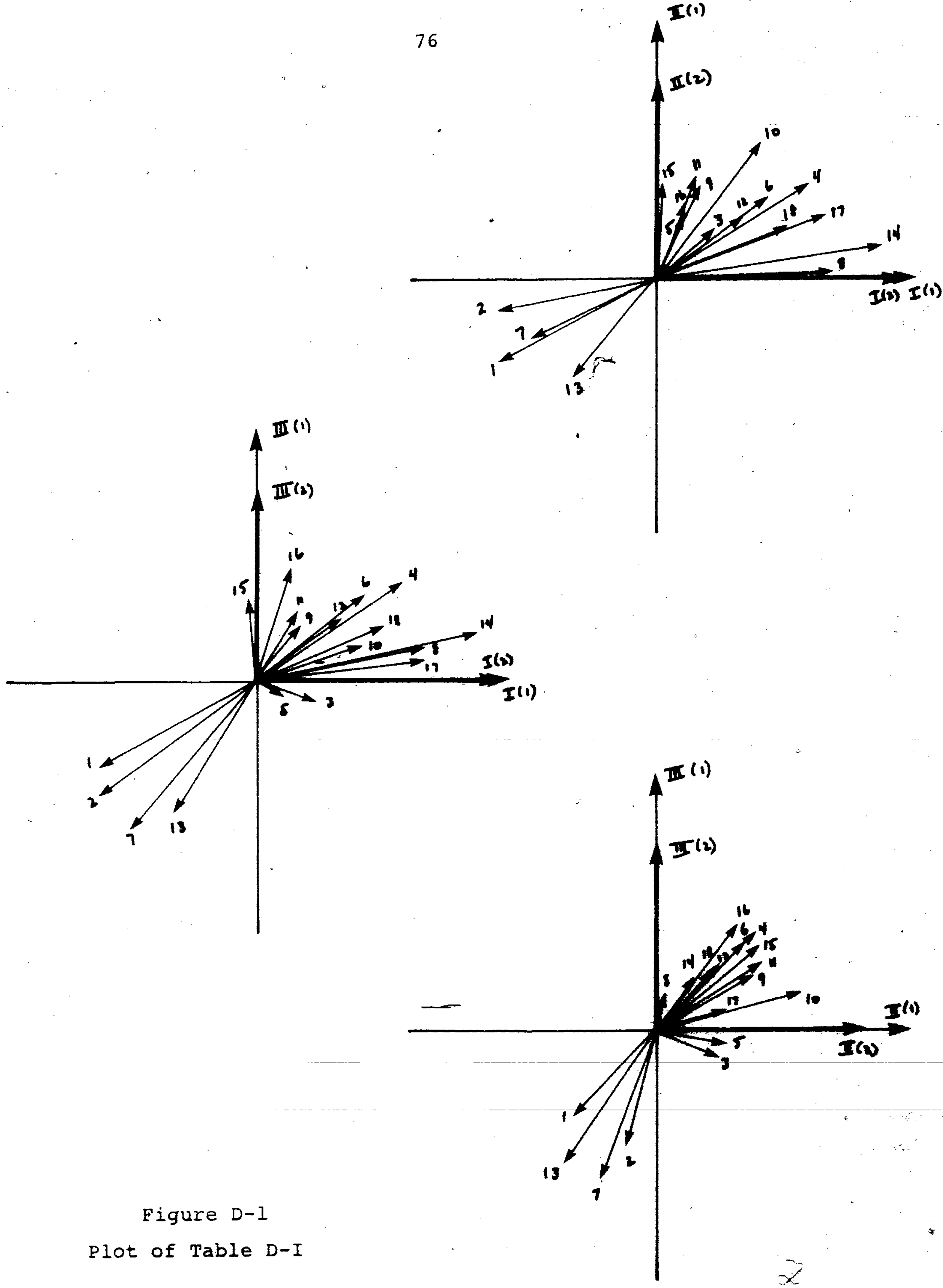


Figure D-1

Plot of Table D-I

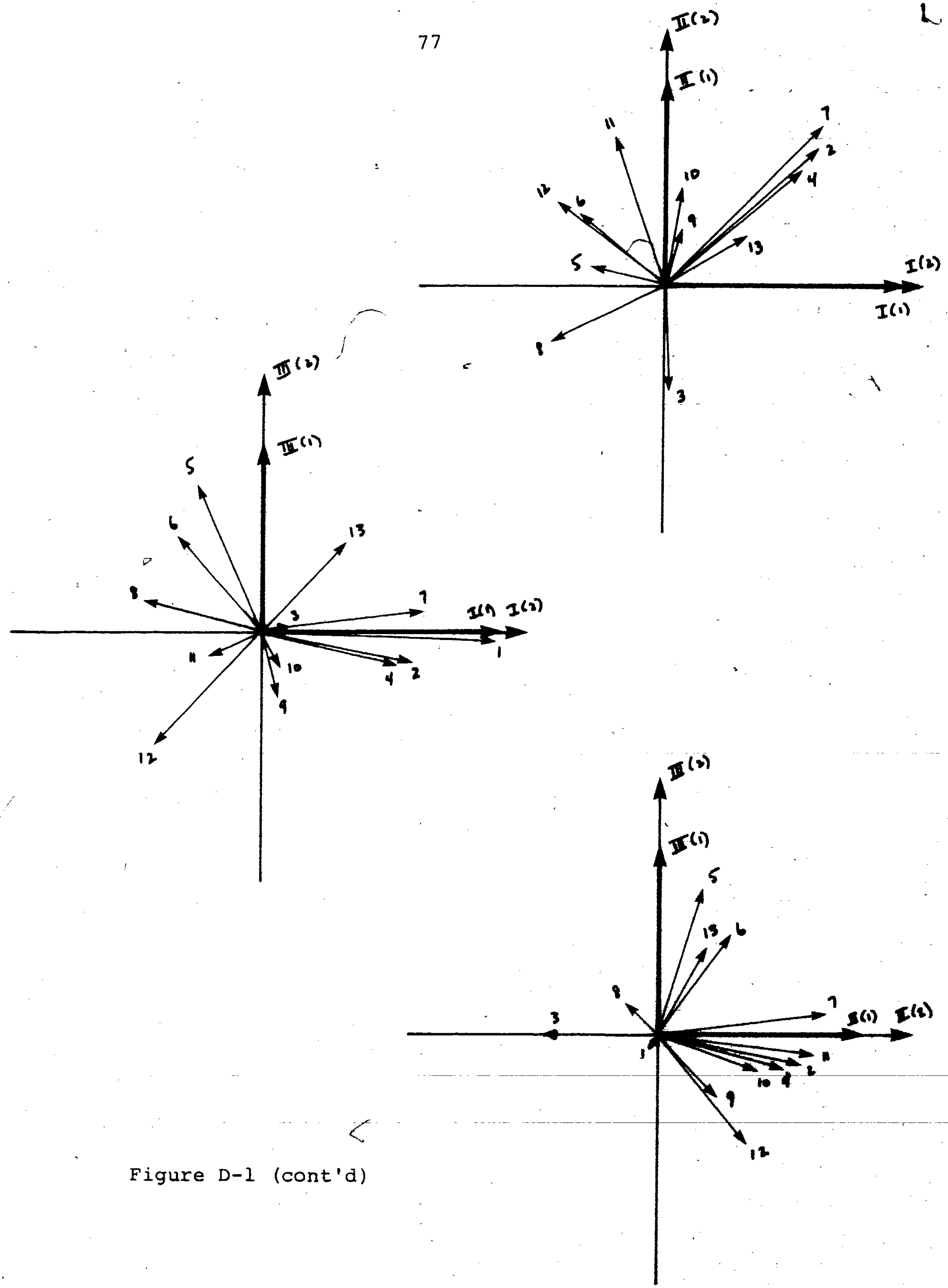


Figure D-1 (cont'd)

Table D-II

Data from Burger, Armentrout, and Rapfogel (1975)

Raw rotation of loadings from canonical analysis of Child's Report of Parental Behavior Inventory and Minnesota Multiphasic Personality Inventory

Observed variables	Canonical loadings		
Set one (parental behavior)	I(1)	II(1)	III(1)
1. Acceptance	-.72 (--)	-.04	-.30 (-)
2. Childcentredness	-.68 (-)	-.19	-.33 (-)
3. Possessiveness	.27	.17	-.03
4. Rejection	.71 (++)	.09	.39 (+)
5. Control	.15	.24	.03
6. Enforcement	.59 (+)	.10	.37 (+)
7. Positive Involvement	-.60 (-)	.12	-.51 (-)
8. Intrusiveness	.69 (+)	-.16	.01
9. Control Through Guilt	.28	.22	.31 (+)
10. Hostile Control	.56 (+)	.37 (+)	.26
11. Inconsistent Discipline	.29	.25	.37 (+)
12. Nonenforcement	.42 (+)	.06	.26
13. Acceptance of Individuation	-.46 (-)	-.08	-.54 (-)
14. Lax Discipline	.92 (++)	-.11	.07
15. Instilling Persistent Anxiety	.16	.23	.44 (+)
16. Hostile Detachment	.24	.10	.48 (+)
17. Withdrawl of Relations	.72 (++)	.09	.05
18. Extreme Autonomy	.58 (+)	.01	.19
Set two (M. M. P. I.)	I(2)	II(2)	III(2)
1. L Scale	.81 (++)	-.31 (-)	-.20
2. F Scale	.77 (++)	.30 (+)	-.03
3. K Scale	-.12	-.37 (-)	-.16
4. Hypochondriasis (Hs)	.68 (+)	.25	-.06
5. Depression (D)	-.18	-.08	.61 (+)
6. Hysteria (Hy)	-.17	.19	.52 (+)
7. Psychopathic Deviate (Pd)	.85 (++)	.29	.18
8. Masculinity-Femininity (Mf)	-.46 (-)	.00	.17
9. Paranoia (Pa)	.15	.28	.15
10. Psychasthenia (Pt)	.22	.37 (+)	.01
11. Schizophrenia (Sc)	.06	.63 (+)	.19
12. Hypomania (Ma)	-.28	.62 (+)	-.19
13. Social Introversion-Extroversion	.41 (+)	-.09	.33 (+)

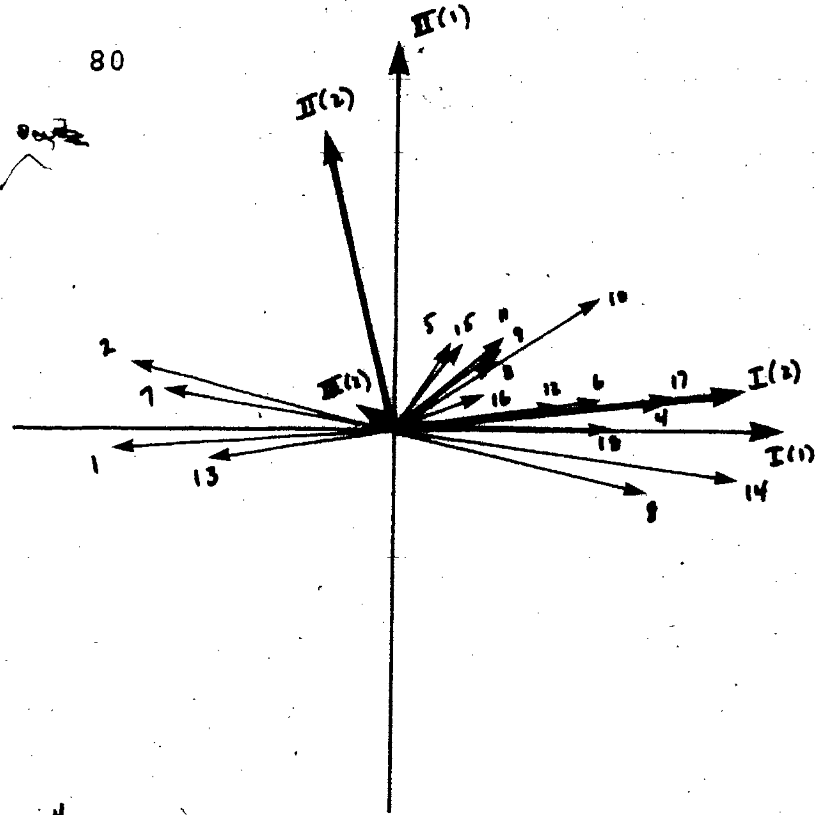
Table D-III

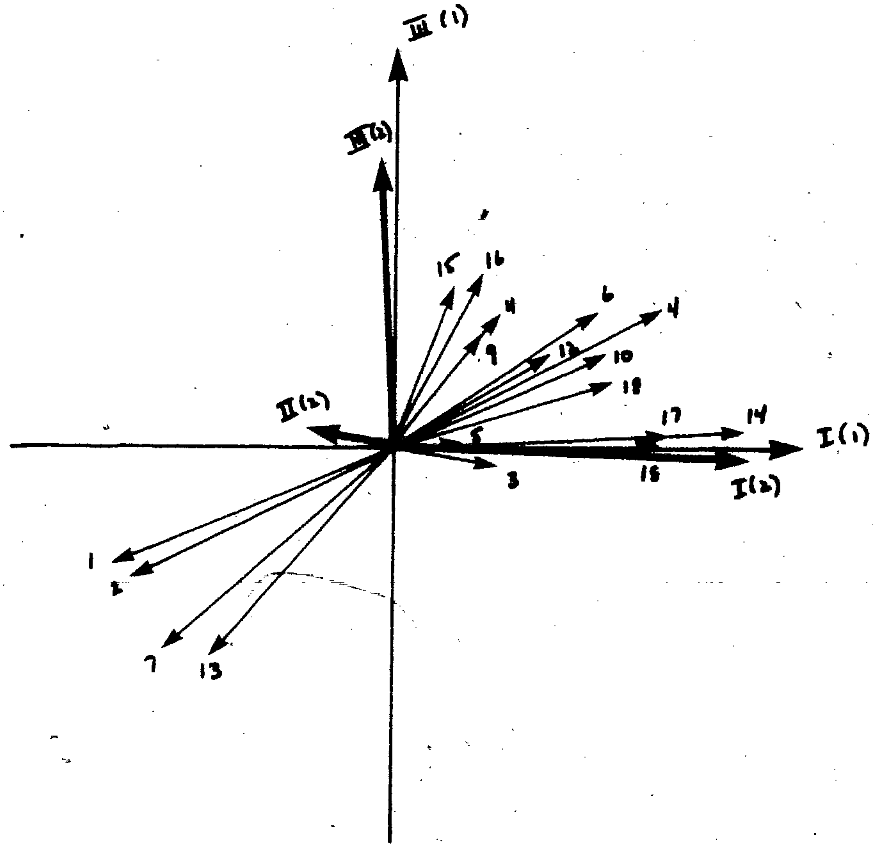
Data from Burger, Armentrout, and Rapfogel (1975)

Variate intercorrelations from canonical analysis of Child's Report of Parental Behavior Inventory and Minnesota Multiphasic Personality Inventory

Raw rotation

		M. M. P. I. variates		
		I(2)	II(2)	III(2)
Parental behavior variates	I(1)	.95	-.19	-.05
	II(1)	.11	.80	.03
	III(1)	-.02	.03	.77





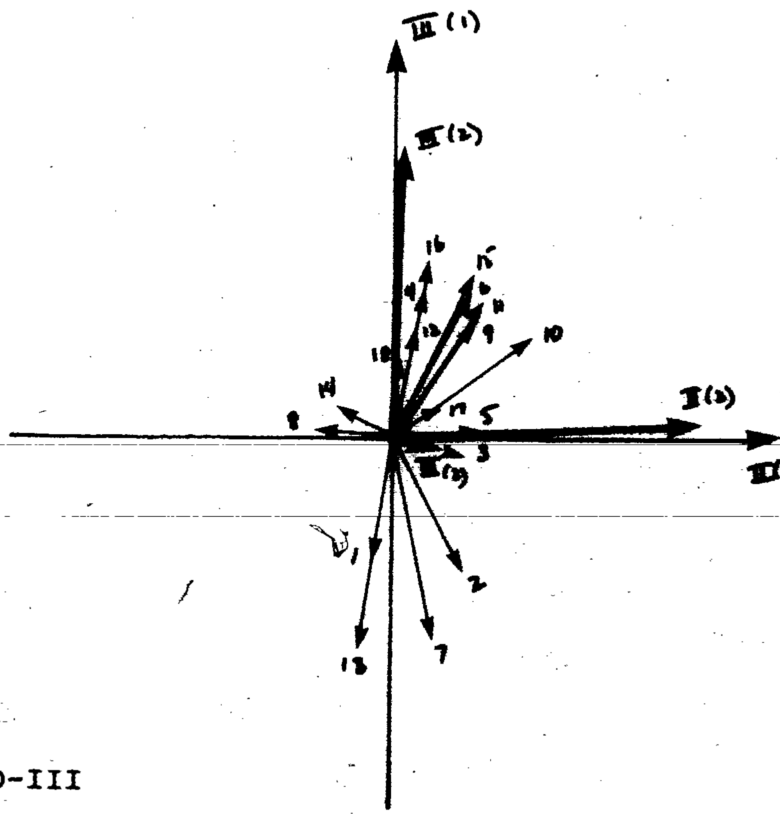


Figure D-2

Plot of Tables D-II and D-III

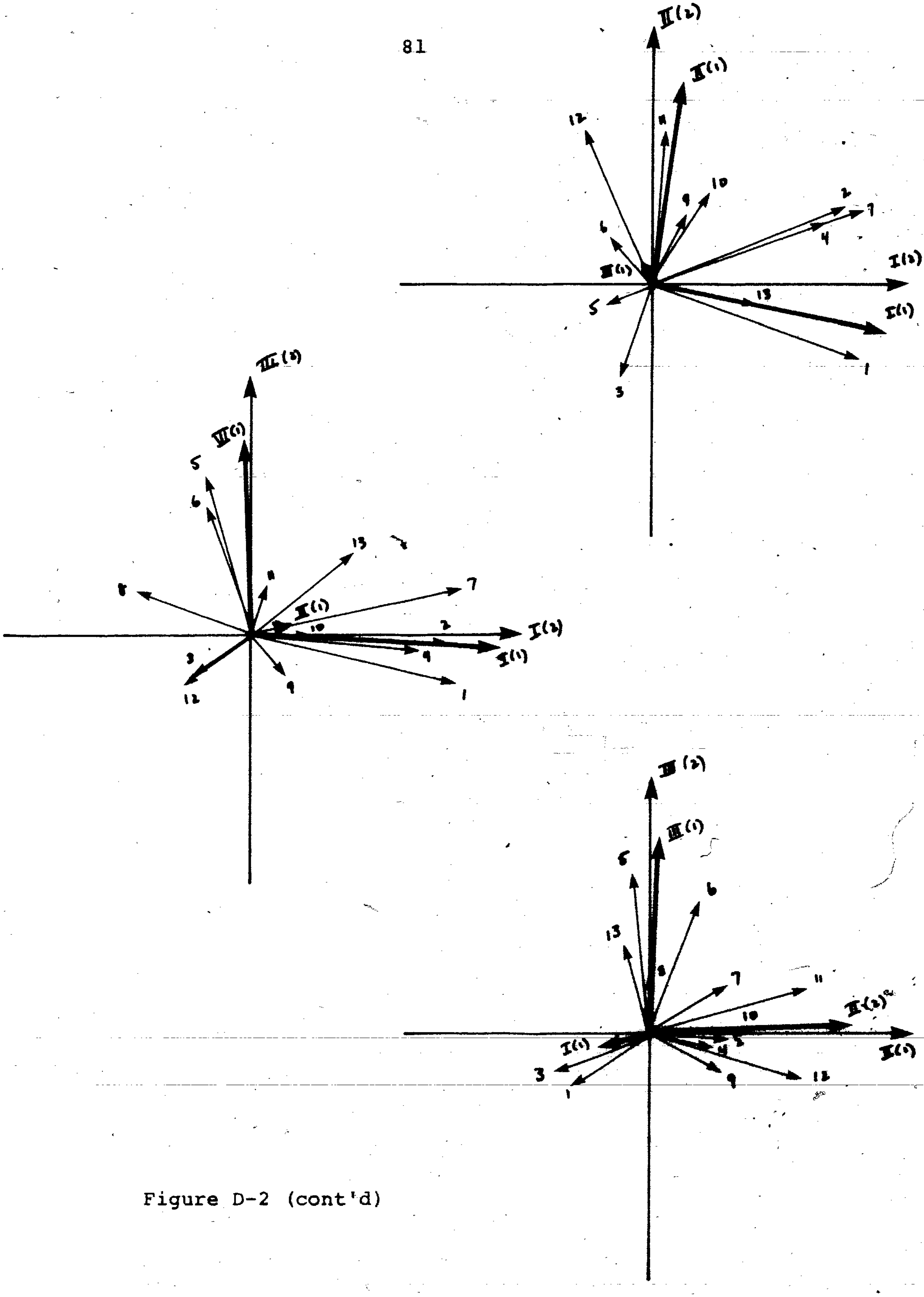


Figure D-2 (cont'd)

Table D-IV

Data from Burger, Armentrout, and Rapfogel (1975)

Normalized rotation of loadings from canonical analysis of Child's Report of Parental Behavior Inventory and Minnesota Multiphasic Personality Inventory

Observed variables	Canonical loadings		
Set one (parental behavior)	I(1)	II(1)	III(1)
1. Acceptance	-.68 (-)	-.17	-.35 (-)
2. Childcentredness	-.69 (-)	.07	-.37 (-)
3. Possessiveness	.22	.23	.00
4. Rejection	.65 (+)	.20	.45 (+)
5. Control	.10	.26	.05
6. Enforcement	.48 (+)	.18	.42 (+)
7. Positive Involvement	-.58 (-)	.03	-.55 (-)
8. Intrusiveness	.70 (++)	-.02	.06
9. Control Through Guilt	.21	.25	.35 (+)
10. Hostile Control	.45 (+)	.45 (+)	.33 (+)
11. Inconsistent Discipline	.21	.27	.40 (+)
12. Nonenforcement	.38 (+)	.13	.29
13. Acceptance of Individuation	-.40 (-)	-.13	-.57 (-)
14. Lax Discipline	.91 (++)	.08	.13
15. Instilling Persistent Anxiety	.08	.22	.46 (+)
16. Hostile Detachment	.18	.11	.50 (+)
17. Withdrawl of Relations	.68 (+)	.23	.11
18. Extreme Autonomy	.55 (+)	.12	.23
Set two (M. M. P. I.)	I(2)	II(2)	III(2)
1. L Scale	.86 (++)	-.12	-.21
2. F Scale	.68 (+)	.47 (+)	-.04
3. K Scale	-.09	-.39 (-)	-.16
4. Hypochondriasis (Hs)	.61 (+)	.39 (+)	-.06
5. Depression (D)	-.15	-.12	.61 (+)
6. Hysteria (Hy)	-.20	.15	.53 (+)
7. Psychopathic Deviate (Pd)	.76 (++)	.47 (+)	.18
8. Masculinity-Femininity (Mf)	-.45 (-)	-.10	.17
9. Paranoia (Pa)	.08	.30 (+)	-.15
10. Psychasthenia (Pt)	.13	.41 (+)	.01
11. Schizophrenia (Sc)	-.08	.62 (+)	.19
12. Hypomania (Ma)	-.41 (-)	.55 (+)	-.18
13. Social Introversion-Extroversion	.42 (+)	.00	.33 (+)

Table D-V

Data from Burger, Armentrout, and Rapfogel (1975)

Variate intercorrelations from canonical analysis of Child's Report of Parental Behavior Inventory and Minnesota Multiphasic Personality Inventory

Normalized rotation

		M. M. P. I. variates		
		I(2)	II(2)	III(2)
Parental behavior variates	I(1)	.96	-.14	-.11
	II(1)	.13	.79	-.04
	III(1)	.05	.07	.76

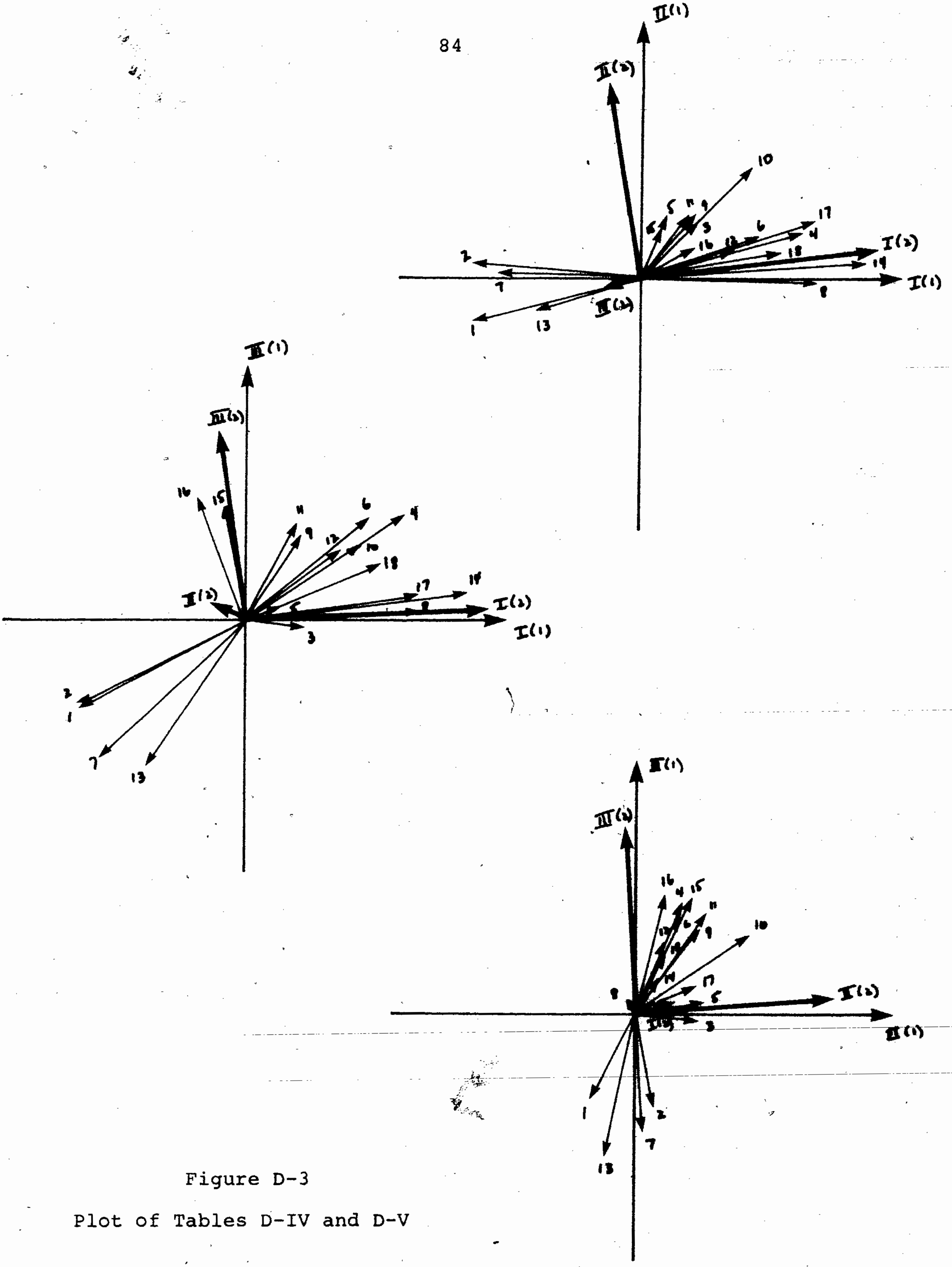
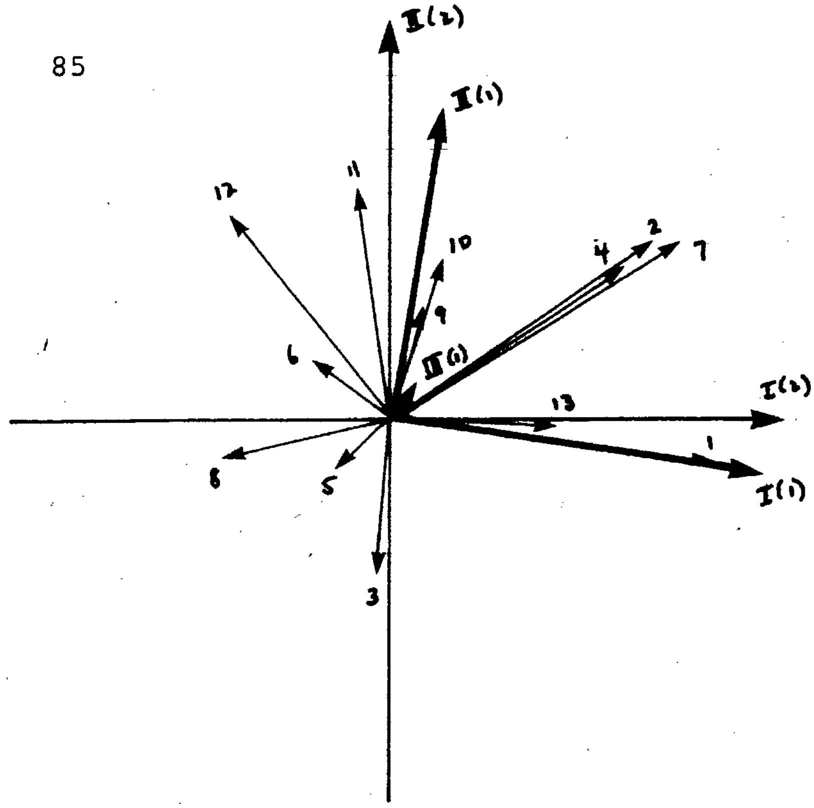
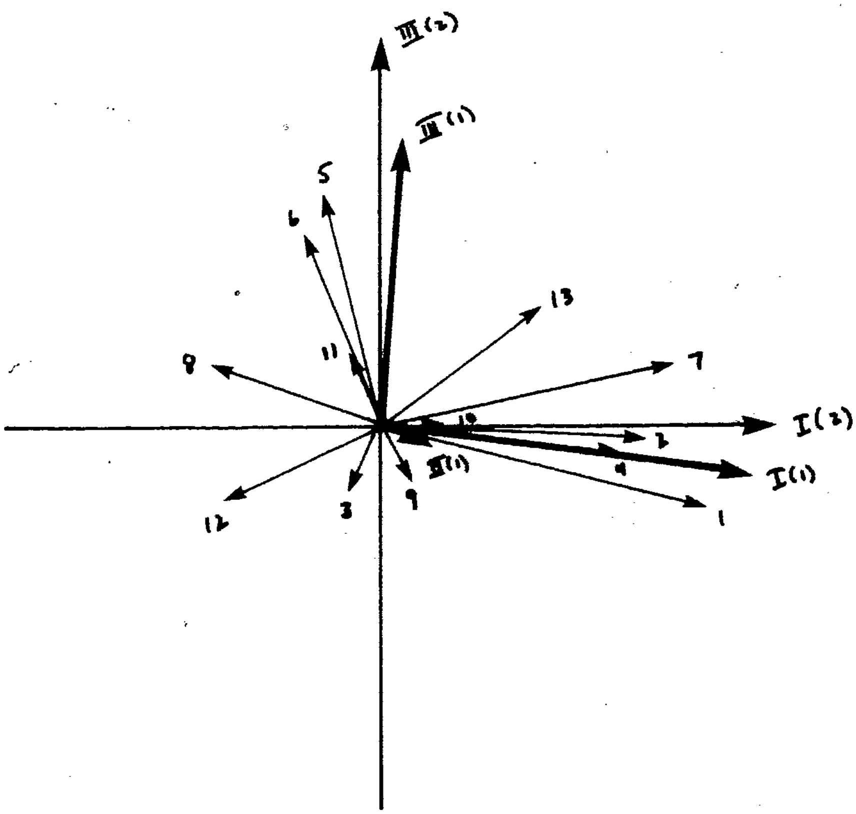


Figure D-3

Plot of Tables D-IV and D-V





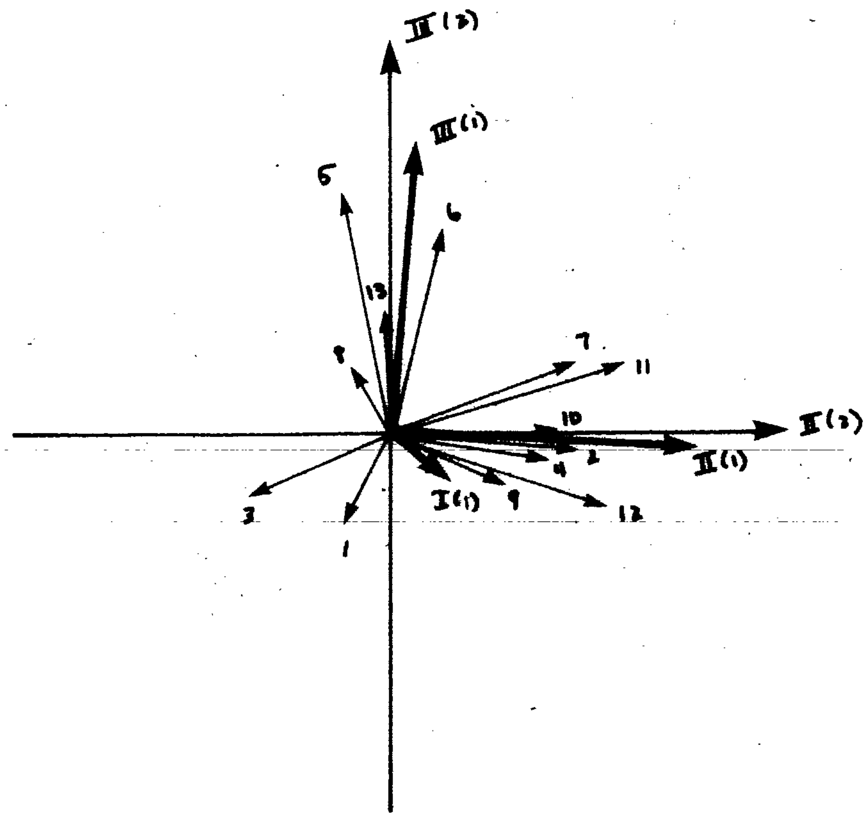


Figure D-3 (cont'd)

Appendix E

Data from Scott (1977)

Crime rates from victimization surveys vs. crime rates from police sources

Table E-I

Data from Scott (1977)

Unrotated loadings from canonical analysis of crime rates from victimization surveys and crime rates from police sources

Observed variables	Canonical loadings		
Set one (victimization rates)	I(1)	II(1)	III(1)
1. Rape	.40 (+)	.25	.51 (+)
2. Robbery	-.80 (--)	-.19	.16
3. Aggravated assault	.42 (+)	.30 (+)	.74 (++)
4. Other assault	.69 (++)	.36 (+)	.40 (+)
5. Burglary	-.15	-.08	.63 (+)
6. Larceny	.58 (+)	.25	.65 (+)
7. Auto theft	-.49 (-)	.77 (++)	.36 (+)
Set two (Police rates)	I(2)	II(2)	III(2)
1. Rape	-.40 (-)	.21	.03
2. Robbery	-.90 (--)	-.23	-.03
3. Aggravated assault	-.48 (-)	-.26	-.37 (-)
4. Burglary	-.15	-.14	.49 (+)
5. Larceny	.38 (+)	-.10	.19
6. Auto theft	-.74 (--)	.62 (+)	.28
7. Murder	-.67 (-)	-.34 (-)	.20
Canonical correlations	.96	.92	.82

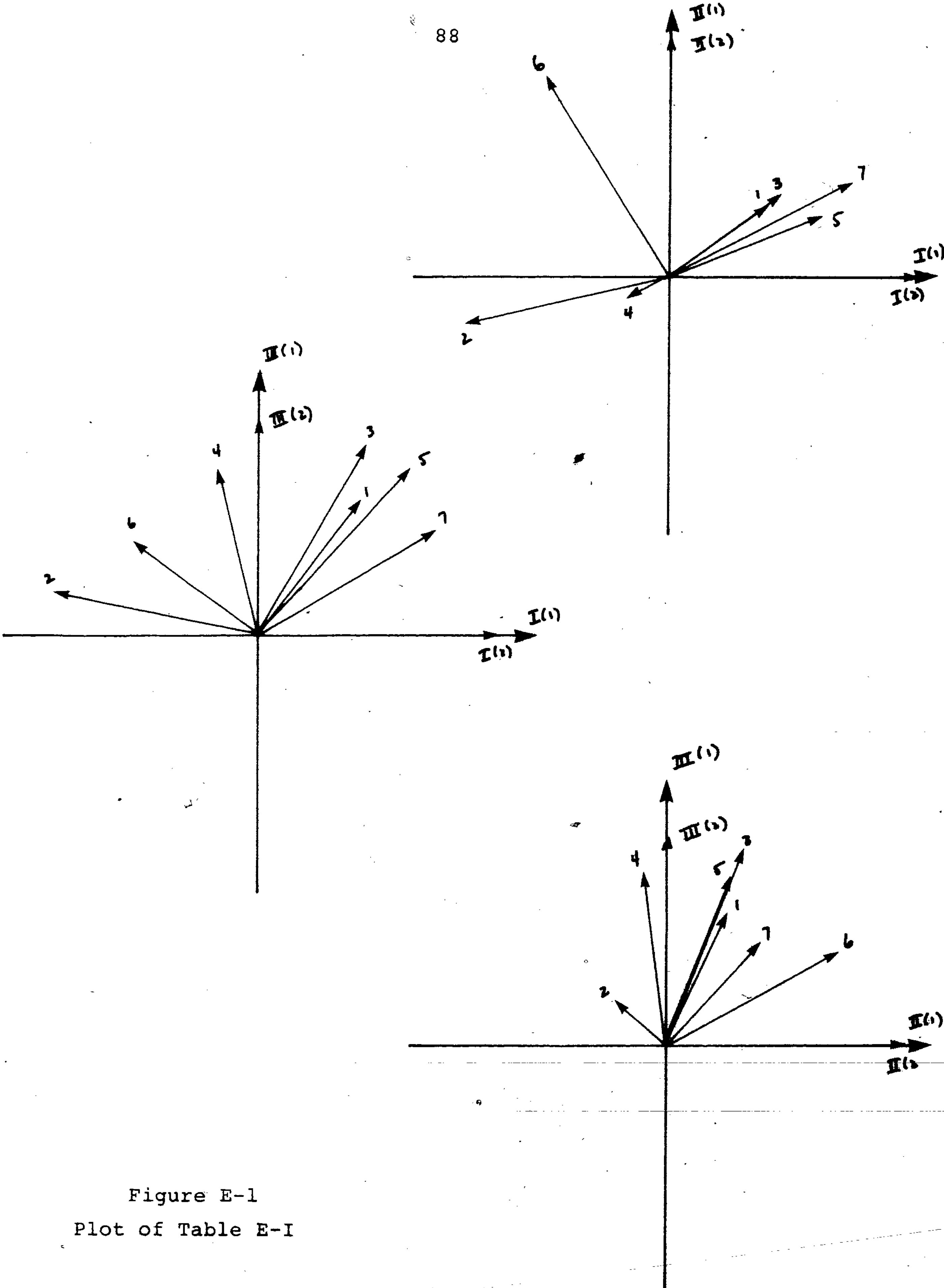


Figure E-1
Plot of Table E-I

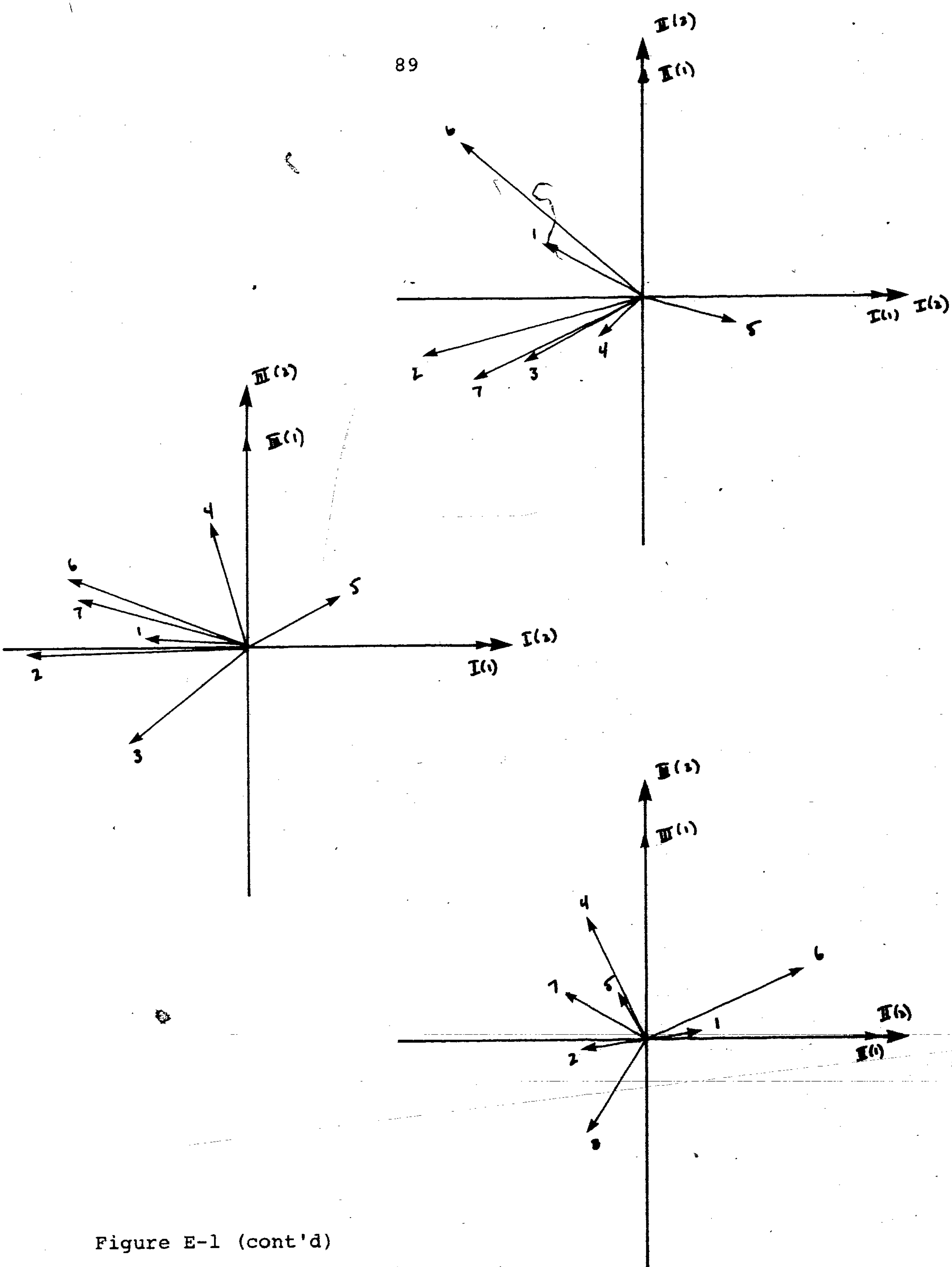


Figure E-1 (cont'd)

Table E-II

Data from Scott (1977)

Raw rotation of loadings from canonical analysis of crime rates from victimization surveys and crime rates from police sources

Observed variables	Canonical loadings		
Set one (victimization rates)	I(1)	II(1)	III(1)
1. Rape	.11	.65 (+)	.22
2. Robbery	.34 (+)	-.63 (-)	.42 (+)
3. Aggravated assault	.20	.78 (++)	.39 (+)
4. Other assault	.01	.88 (++)	.01
5. Burglary	.19	.13	.62 (+)
6. Larceny	.04	.86 (++)	.30 (+)
7. Auto theft	.97 (++)	.09	.07
Set two (Police rates)	I(2)	II(2)	III(2)
1. Rape	.40 (+)	-.21	-.02
2. Robbery	.33 (+)	-.85 (--)	.16
3. Aggravated assault	-.02	-.63 (-)	-.19
4. Burglary	.12	-.04	.52 (+)
5. Larceny	-.25	.32 (+)	.18
6. Auto theft	.98 (++)	-.21	.08
7. Murder	.18	-.65 (-)	.38 (+)

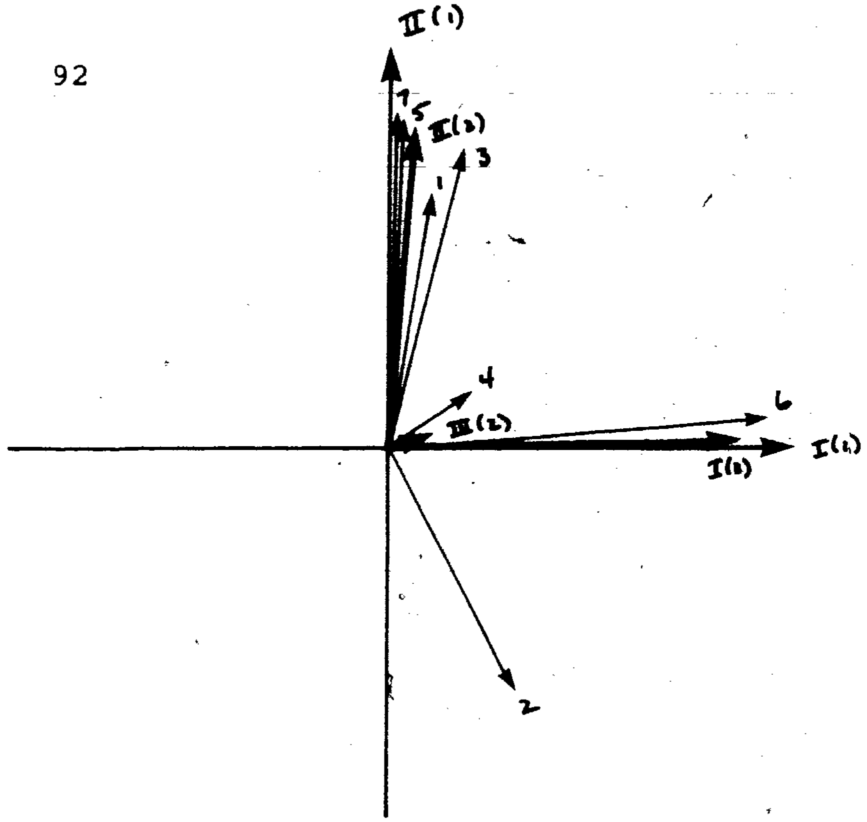
Table E-III

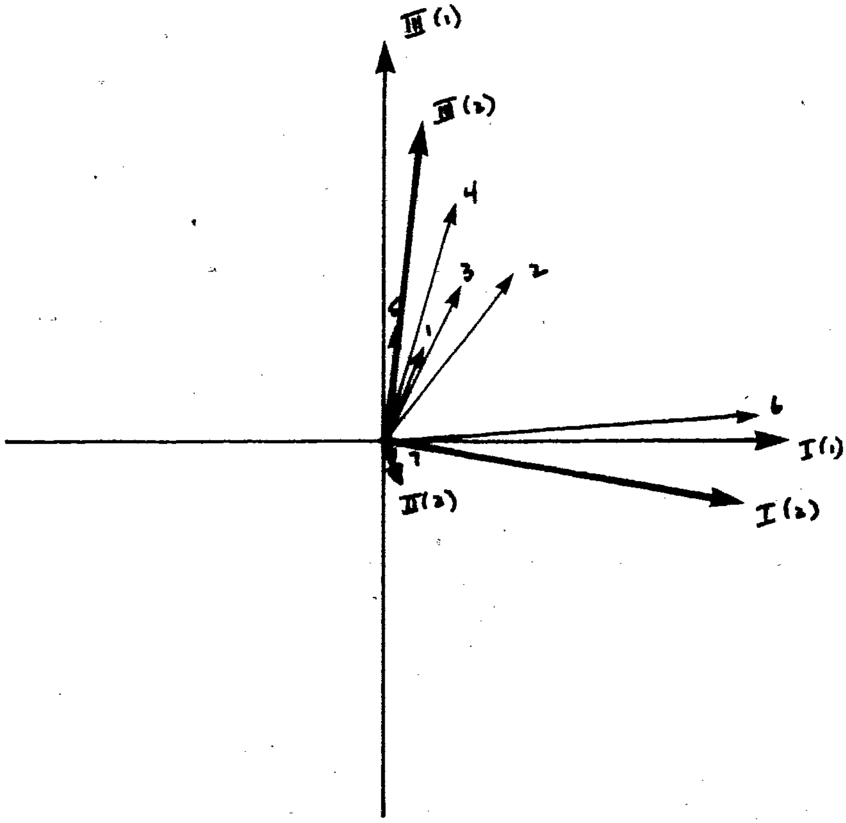
Data from Scott (1977)

Variate intercorrelations from canonical analysis of crime rates from victimization surveys and crime rates from police sources

Raw rotation

		Police variates		
		I(2)	II(2)	III(2)
Victimization variates	I(1)	.92	.01	.04
	II(1)	.05	.92	.10
	III(1)	.00	-.18	.83





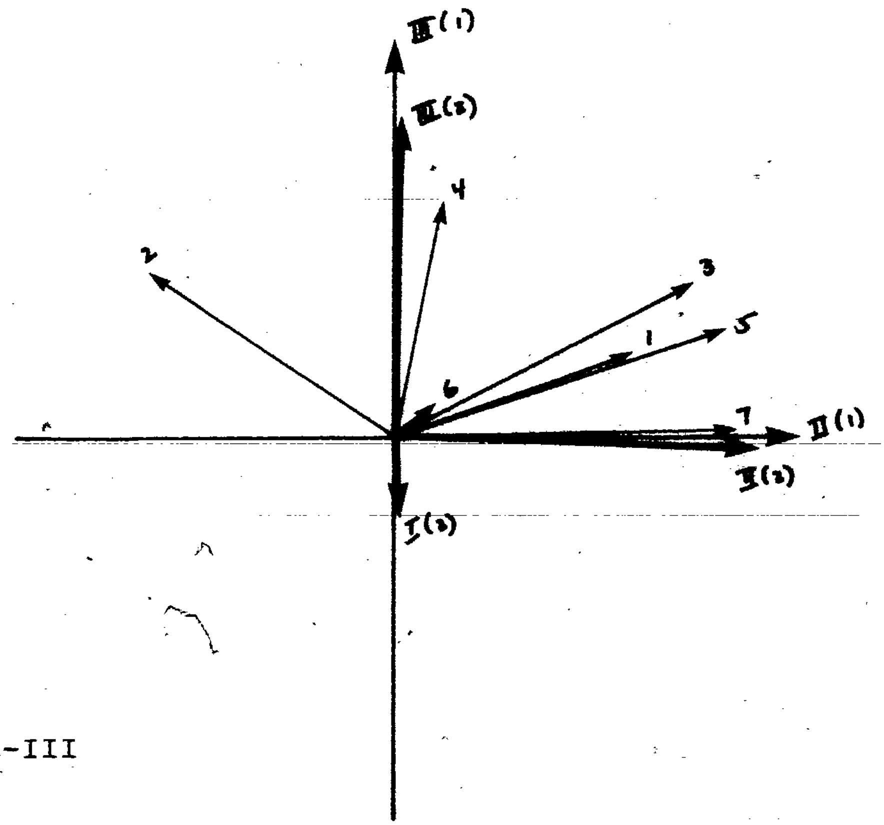


Figure E-2

Plot of Tables E-II and E-III

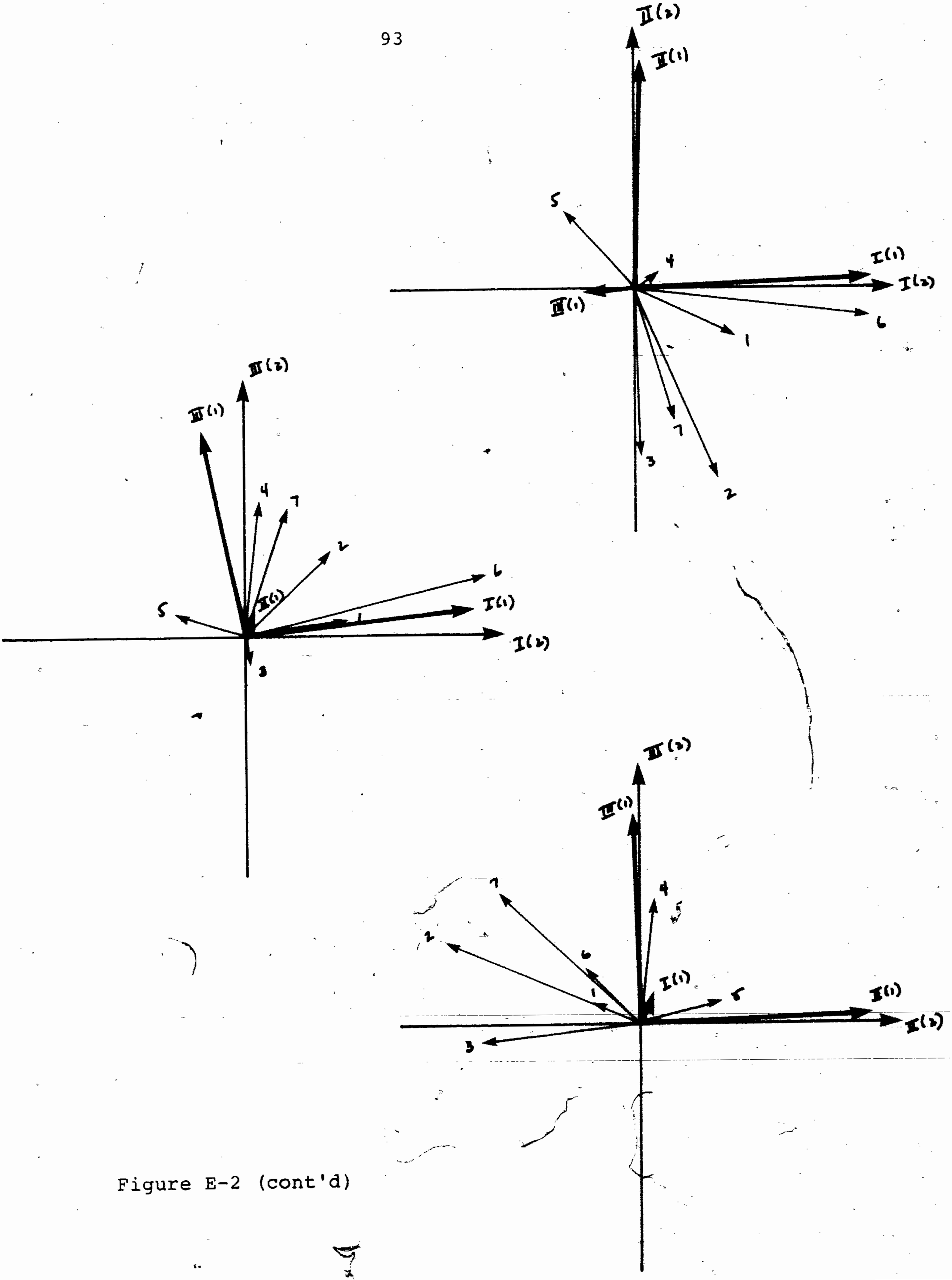


Figure E-2 (cont'd)

Table E-IV

Data from Scott (1977)

Normalized rotation of loadings from canonical analysis of crime rates from victimization surveys and crime rates from police sources

Observed variables	Canonical loadings		
Set one (victimization rates)	I (1)	II (1)	III (1)
1. Rape	.07	.66 (+)	.19
2. Robbery	.32 (+)	-.59 (-)	.49 (+)
3. Aggravated assault	.14	.81 (++)	.36 (+)
4. Other assault	-.02	.88 (++)	-.05
5. Burglary	.13	.17	.62 (+)
6. Larceny	-.01	.87 (++)	.24
7. Auto theft	.96 (++)	.13	.15
Set two (Police rates)	I (2)	II (2)	III (2)
1. Rape	.41 (+)	-.18	.06
2. Robbery	.35 (+)	-.79 (--)	.34 (+)
3. Aggravated assault	.03	-.65 (-)	-.08
4. Burglary	.06	.06	.53 (+)
5. Larceny	-.28	.32 (+)	.09
6. Auto theft	.97 (++)	-.12	.22
7. Murder	.17	-.56 (-)	.51 (+)

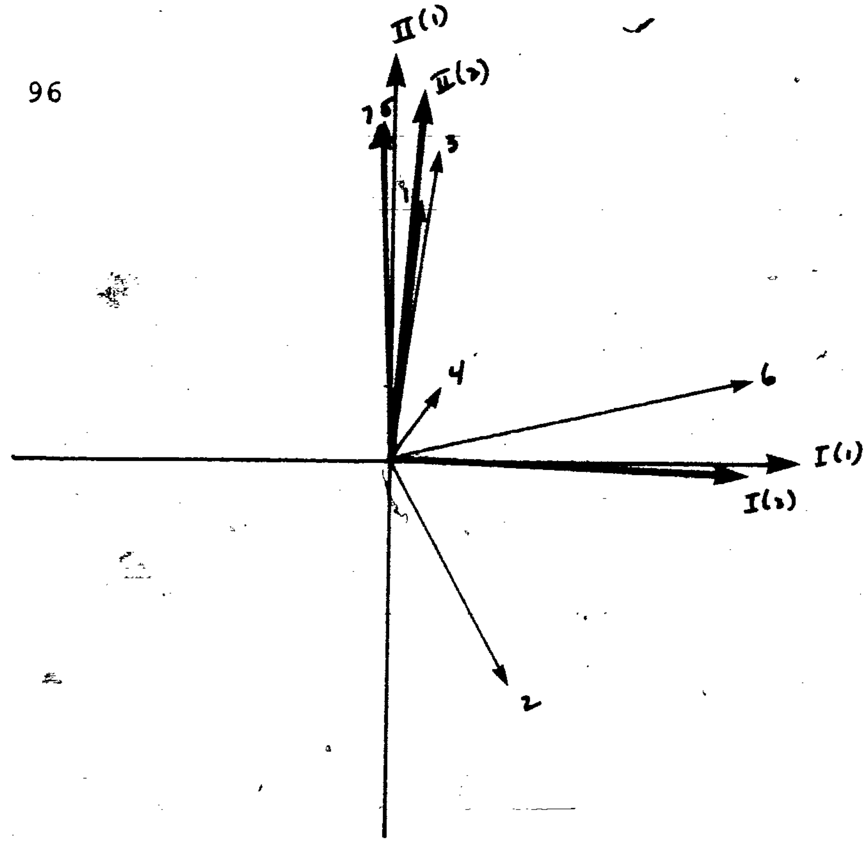
Table E-V

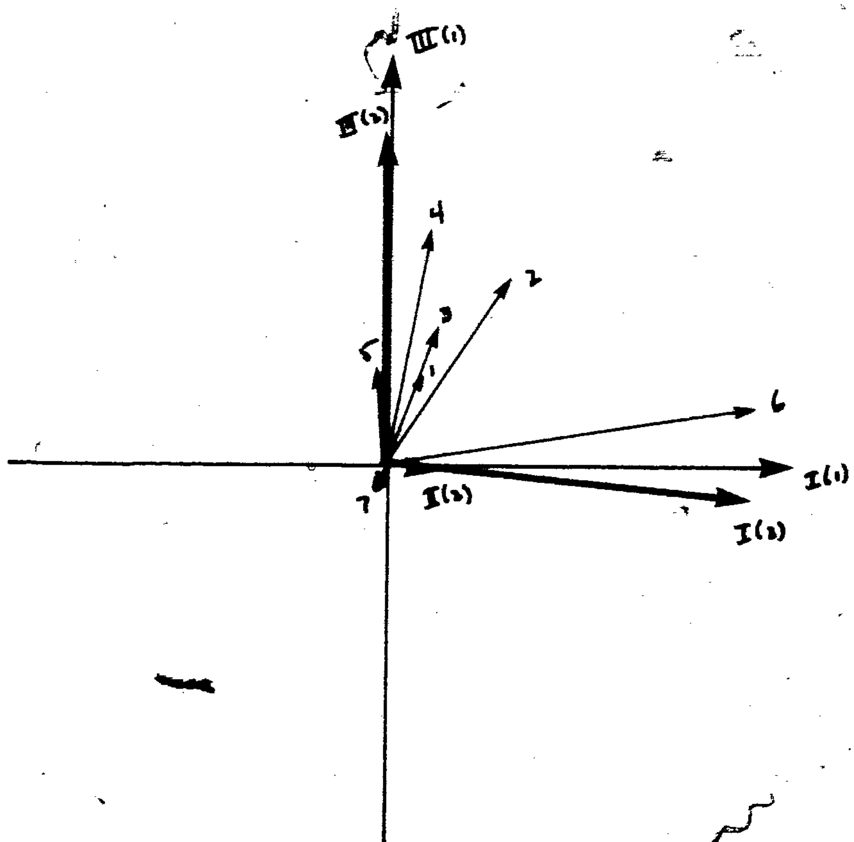
Data from Scott (1977)

Variate intercorrelations from canonical analysis of crime rates from victimization surveys and crime rates from police sources

Normalized rotation

		Police variates		
		I(2)	II(2)	III(2)
Victimization variates	I(1)	.93	-.03	.02
	II(1)	.08	.91	-.01
	III(1)	-.01	-.10	.85





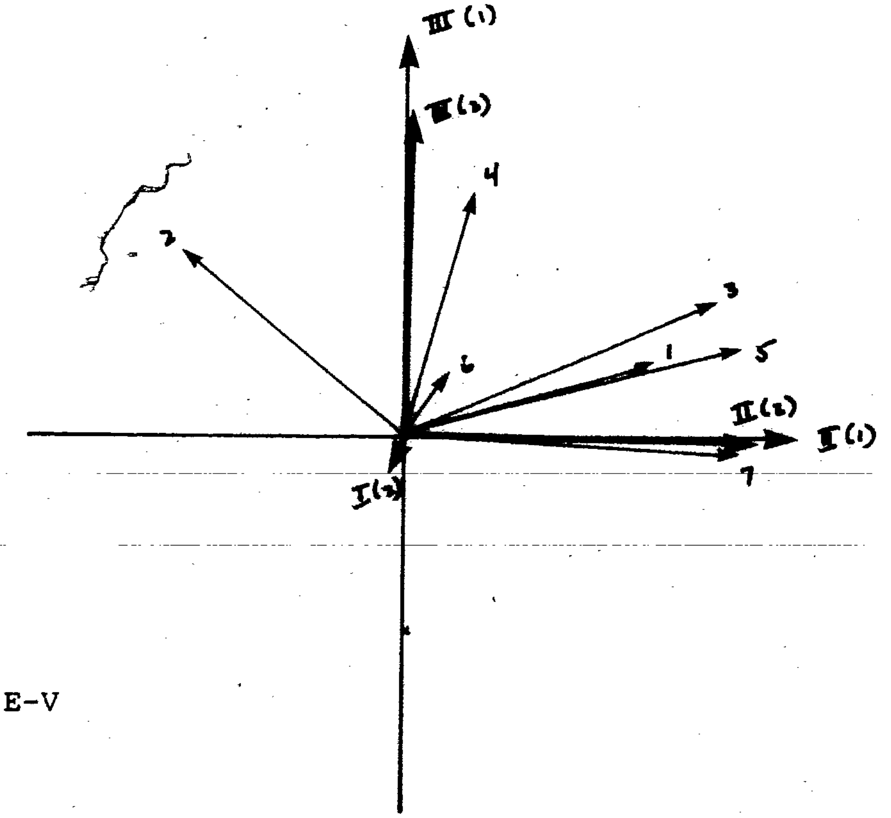
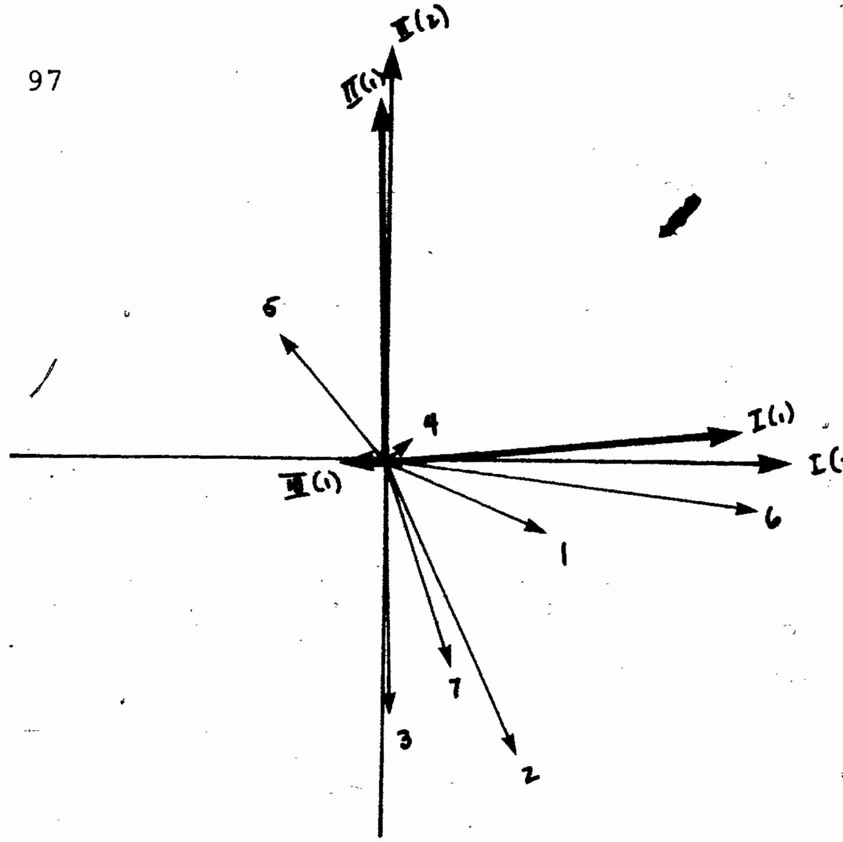
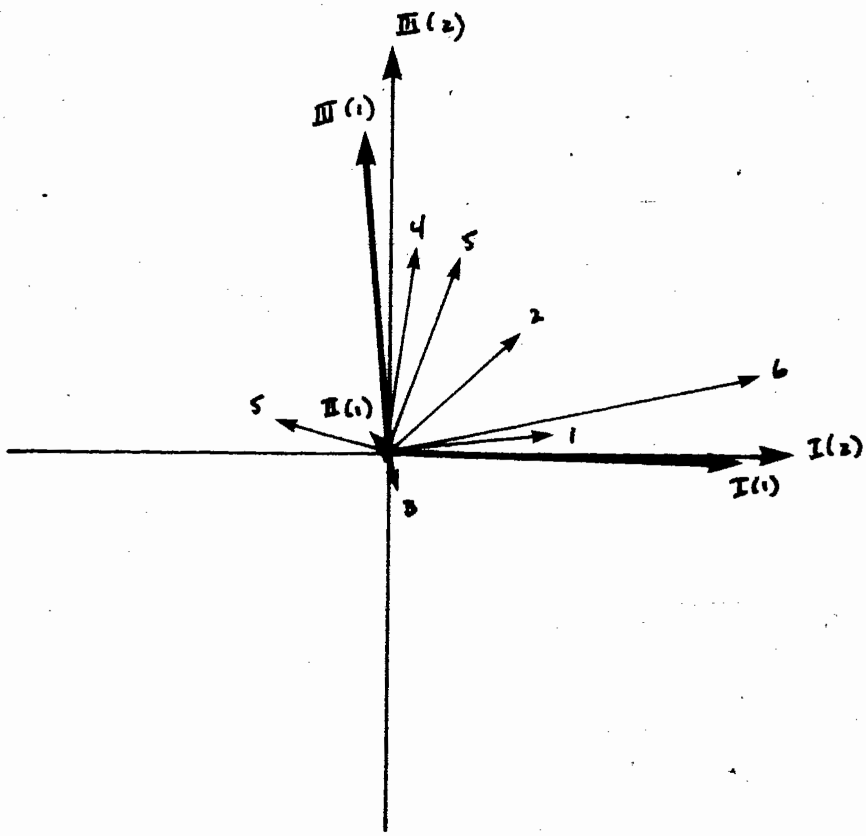


Figure E-3

Plot of Tables E-IV and E-V





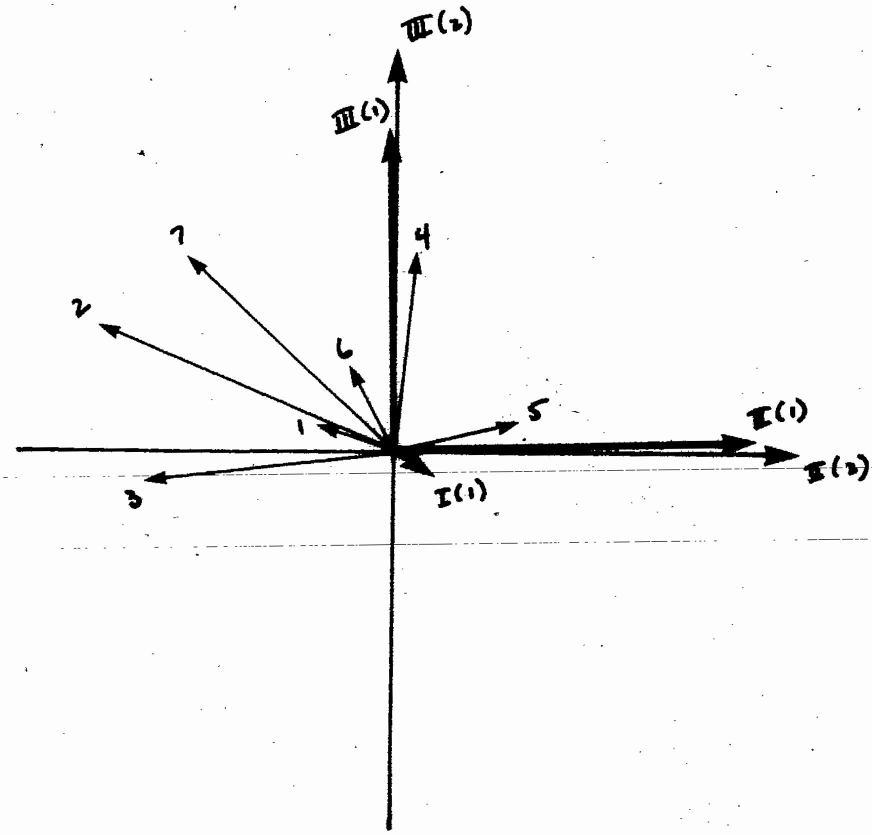


Figure E-3 (cont'd)

REFERENCES

Anderson, T.W. An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Bartlett, M.S. The statistical significance of canonical correlation. Biometrika, 1941, 32, 29-38.

Bartlett, M.S. Multivariate analysis. Journal of the Royal Statistical Society, Series B, 1947, 9, 176-197.

Bock, R.D. Multivariate Statistical Methods in Behavioral Research. New York: McGraw-Hill, 1975.

Burger, G.K., Armentrout, J.A., and Rapfogel, R.G. Recalled parental behavior and objective personality measures: a canonical analysis. Journal of Personality Assessment, 39, 514-522.

Carroll, J.D. Generalization of canonical correlation to three or more sets of variables. Proceedings, American Psychological Association, 76th Annual Convention, 1968, pp. 227-28.

Cattell, R.B., Eber, H.W., and Tatsuoka, M.M. Handbook for the Sixteen Personality Factor Questionnaire. Champaign, Ill.: Institute for Personality and Ability Testing, 1970.

Cliff, N. and Krus, D.J. Interpretation of canonical analysis: rotated vs. unrotated solutions. Psychometrika, 1976, 41, 35-42.

Cooley, W.W. and Lohnes, P.R. Multivariate Data Analysis. New York: Wiley, 1971.

Dempster, A.P. Elements of Continuous Multivariate Analysis. Reading, Mass.: Addison-Wesley, 1969.

Dykstra, R. Summary of the second-grade phase of the Cooperative Research Program in Primary Reading Instruction. Reading Research Quarterly, 1968, 4, 49-70.

Hakstian, A.R. Two-matrix orthogonal rotation procedures. Psychometrika, 1976, 41, 267-272.

Harman, H.H. Modern Factor Analysis. (2nd. ed.) Chicago: University of Chicago Press, 1967.

Harris, C.W. and Kaiser, H.F. Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 1964, 29, 347-362.

Harris, R.J. A Primer of Multivariate Statistics. New York: Academic Press, 1975.

Hathaway, S.R. and McKinley, J.C. Manual for the Minnesota Multiphasic Personality Inventory. (Rev. ed.) New York: Psychological Corporation, 1951.

Holland, J.L. Manual for the Vocational Preference Inventory. (6th rev.) Iowa City: Educational Research Associates, 1965.

Hotelling, H. The most predictable criterion. Journal of Educational Psychology, 1935, 26, 139-142.

Hotelling, H. Relations between two sets of variates. Biometrika, 1936, 28, 321-377.

Kaiser, H.F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Kendall, M.G. A Course In Multivariate Statistics. London: Griffin, 1961.

Lawley, D.N. and Maxwell, A.E. Factor Analysis as a Statistical Method. (2nd ed.) London: Butterworths, 1971.

Lohnes, P.R. and Gray, M.M. Intelligence and the cooperative reading studies. Reading Research Quarterly, 1972, 7, 466-476.

Lohnes, P.R. and Marshall, T.O. Redundancy in student records. American Educational Research Journal, 1965, 2, 19-23.

Miller, J.K. In defense of the general canonical correlation index: reply to Nicewander and Wood. Psychological Bulletin, 1975, 82, 207-209.

Miller, J.K. and Farr, S.D. Bi-multivariate redundancy: a comprehensive measure of interbattery relationship. Multivariate Behavioral Research, 1971, 6, 313-324.

Morrison, D.F. Multivariate Statistical Methods. (2nd ed.) New York: McGraw-Hill, 1976.

Mulaik, S.A. The Foundations of Factor Analysis. New York; McGraw-Hill, 1972.

Press, S.J. Applied Multivariate Analysis. New York: Holt, Rinehart, and Winston, 1972.

Rao, C.R. Linear Statistical Inference and its Applications. New York: Wiley, 1965.

Schaefer, E.S. A configurational analysis of children's reports of parent behavior. Journal of Consulting Psychology, 1965, 29, 552-557.

Scott, D.W. Components of crime measured two ways: a comparison of victimization rates and police rates. Unpublished manuscript, Department of Criminology, Simon Fraser University, Burnaby, British Columbia, 1977.

Stewart, D. and Love, W. A general canonical correlation index. Psychological Bulletin, 1968, 70, 160-163.

Tatsuoka, M.M. Multivariate Analysis: Techniques for Educational and Psychological Research. New York: Wiley, 1971.

Thurstone, L.L. Multiple Factor Analysis. Chicago: University of Chicago Press, 1947.

Van De Geer, J.P. Introduction to Multivariate Analysis for the Social Sciences. San Francisco: Freeman, 1971.

Wilks, S.S. Certain generalizations in the analysis of variance. Biometrika, 1932, 24, 471-494.

Williams, J.D. and Williams, C.M. Canonical analysis of the Vocational Preference Inventory and the Sixteen Personality Factor Questionnaire. Psychological Reports, 1973, 32, 211-214.