DOCUMENT RESUME

ED 282 904 TM 870 302

AUTHOR Thompson, Bruce
TITLE Fundamentals of Canonical Correlation Analysis: Basics and Three Common Fallacies in Interpretation.
PUB DATE 16 Apr 87
NOTE 24p.; Paper presented at the Annual Meeting of the Society for Multivariate Experimental Psychology, Southwestern Division (New Orleans, LA, April 16, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Correlation; *Multivariate Analysis; *Research Problems; Statistical Bias; *Statistical Inference; Statistical Studies

ABSTRACT

Canonical correlation analysis is illustrated and three common fallacious interpretation practices are described. Simply, canonical correlation is an example of the bivariate case. Like all parametric methods, it involves the creation of synthetic scores for each person. It presumes at least two predictor variables and at least two criterion variables. Weights, usually labelled standardized function coefficients, are applied to each individual's data to yield the synthetic variables which are the basis for canonical analysis. However, in canonical correlation, several sets of weights and synthetic variables can be created. Structure coefficients and index coefficients may also be computed. The interpretation of canonical results is challenging because of the myriad of coefficients produced. The three common fallacies to be avoided are: (1) interpreting structure coefficients while ignoring function coefficients; (2) interpreting redundancy coefficients; and (3) failing to employ commonality analysis. When researchers are aware of these pitfalls, canonical correlation analysis can be a powerful analytic method that may be the best technique in a complex situation. (GDC)

swsmep.rno 4/13/87 11:30 am

FUNDAMENTALS OF CANONICAL CORRELATION ANALYSIS: BASICS AND THREE COMMON FALLACIES IN INTERPRETATION

Bruce Thompson

University of New Orleans 70148

and

Louisiana State University Medical Center

Paper presented as invited Feature Speaker at the annual meeting of the Society for Multivariate Experimental Psychology--Southwestern Division, New Orleans, April 16, 1987.

ABSTRACT

The paper briefly explains the logic underlying the basic calculations employed in canonical correlation analysis. The paper also discusses three common fallacious interpretation practices that may lead to incorrect conclusions based on canonical results. A small hypothetical data set is employed to make the discussion concrete. It is suggested that canonical correlation analysis is a powerful analytic method that frequently best honors the complex nature of the reality about which the researcher wishes to generalize.

Several trends in analytic practice seem to be discernable as incremental changes that are moving social science slowly toward more productive inquiry. For example, researchers have increasingly recognized that statistical significance may not be a particularly effective criterion with which to evaluate results (Thompson, 1987); popular developments in meta-analysis (Jones & Fiske, 1953; Glass, McGaw & Smith, 1981; Rosenthal, 1984) may have compelled more researchers to recognize the importance of effect sizes in their studies. Researchers have also increasingly recognized that statistical control, such as that employed in analysis of covariance (ANCOVA), must be used with extraordinary caution; these methods tend to either be unnecessary or seriously distort results (Thompson, 1986b, pp. 19-25) and lead to "tragically misleading analyses" (Campbell & Erlebacher, 1975, p. 597).

However, the trend away from the use of classical analysis of variance methods (Goodwin & Goodwin, 1985) may be the most noteworthy trend of all, since the use of the methods can have several deleterious effects (Cohen, 1968; Thompson, 1986a). Even when analysis of variance methods represent good analytic choices, regression or general linear model approaches to the methods still tend to be superior since these approaches tend to yield greater power against Type II error and tend to be more theoretically grounded (Thompson, 1985a).

The gradual shift away from the use of analysis of variance approaches has been due in part to an increased recognition that all parametric univariate methods are special cases of regression analysis (Cohen, 1968). The shift has also been due to increased

recognition that many researchers

> prefer experimental over correlational research designs because experimental designs provide more complete information about causality. Why does this situation contribute to OVAism? Because some researchers confuse research designs with the statistical techniques which are used to analyze the data which the designs help to generate. (Thompson, 1981, p. 5)

As Thompson (1986b, p. 17) notes,

> The fact that OVA methods are appropriate when predictor variables such as experimental assignment naturally occur at the nominal level of scale has stimulated some researchers to unconsciously [and incorrectly] associate the consequences of experimental design selection with OVA methods.

However, in reality all parametric significance tests, including those which are multivariate, are special cases of canonical correlation analysis (Knapp, 1978). Indeed, Thompson (1985b) illustrates how various univariate and multivariate analyses can all be conducted using canonical correlation analysis. Thompson (1986c) notes that the evaluation of several hypothesis tests within a single study inflates the experimentwise Type I error probability, usually to a somewhat unknown degree. The failure to use multivariate methods often also distorts the reality about which the researcher is attempting to generalize--the least of these distortions occurs

when a researcher completes several univariate tests and finds no statistically significant results when significance would have occurred if a multivariate test had been employed (Thompson, 1986c). Thompson (1986c) presents a data set illustrating how this can occur. These various considerations suggest that canonical correlation analysis may be a powerful and important weapon in the social scientist's arsenal of analytic weapons.

The purpose of the present paper is to briefly explain the logic underlying the basic calculations employed in canonical correlation analysis. The paper also discusses three common fallacious interpretation practices that may lead to incorrect conclusions based on canonical results. A small hypothetical data set is employed to make the discussion concrete.

The Basic Logic of Canonical Calculations

Thompson (1983) notes that canonical correlation can be considered as an example of the bivariate case. This conceptualization has instructional appeal because most students feel comfortable working with bivariate correlation coefficients. The view is also important because it forces realization that canonical analysis, like all parametric methods, involves the creation of "synthetic" scores for each person. In regression analyses the synthetic scores are the predicted dependent variable scores of each of the subjects, sometimes termed "YHAT"; the correlation between the subjects' actual and synthetic variables is the multiple correlation coefficient, while the sum of squares of the "YHAT" scores equals the sum of squares explained. In factor analysis these synthetic variables are the

factor scores of each subject on each of the factors. In discriminant analysis these synthetic variables are the discriminant scores of each subject on each of the discriminant functions.

Table 1 presents a hypothetical data set that will be employed to illustrate how scores of individuals are converted into the synthetic variables that are actually the focus of a canonical correlation analysis. The data are adapted from those presented by Harris (1987). The data set involves two criterion variables, "X" and "Y," and two predictor variables, "A" and "B". Since canonical correlation analysis presumes at least two predictor and at least two criterion variables, the data set represents the simplest case for which a true canonical analysis can be conducted. If a canonical analysis of a smaller data set was conducted, most researchers would refer to the analysis using some other name, such as multiple regression analysis. Table 1 also presents each of the five persons' scores on the four variables converted into their equivalent Z-score forms.

INSERT TABLE 1 ABOUT HERE.

Various analytic methods yield weights that are applied to variables to optimize some condition--such weights include beta weights, factor pattern coefficients, and discriminant function coefficients. These weights are all equivalent, but in canonical correlation analysis the weights are usually labelled standardized function coefficients. These weights are applied to each individual's data to yield the synthetic variables that are the basis for canonical analysis.

However, in canonical analysis several sets of weights and of

the resulting synthetic variables can be created. These canonical functions are related to factors, are uncorrelated or orthogonal, and can be rotated in various ways (Thompson, 1984). The number of functions that can be computed in a canonical analysis equals the number of variables in the smaller of the two variable sets, as explained by Thompson (1984). In the present example, since each variable set consisted of two variables, two canonical functions could be computed. Table 2 presents the canonical function coefficients and other selected results from the analysis.

<u>INSERT TABLE 2 ABOUT HERE.</u>

Table 3 illustrates the computation of the synthetic variables for each of the five subjects using the Function I function coefficients; the reader may wish to compute the corresponding values associated with the Function II results. For a given function, two synthetic scores are produced for each subject--one associated with the composite of weighted criterion variables, and one associated with the composite of weighted predictor variables. For example, as noted in Table 3, the criterion synthetic variable score, "CRITCOMP," for subject one was 1.29589 ([-1.44986*-1.35287] + [+1.04101*-.63850]). By the same token, the predictor synthetic variable score for subject five was -1.21913 ([-1.58021*+1.32563] + [1.24215*+.67606]).

<u>INSERT TABLE 3 ABOUT HERE</u>

The canonical correlation (<u>R</u>c) is nothing more (or less) than the Pearson product-moment correlation between the synthetic variable scores of the subjects on a given function. This can be

illustrated in several ways using the present results. For example, for this case, the bivariate correlation equals the sum of the cross-products of the two variables, the sum then being divided by $n - 1$. The cross products of the synthetic variables for each of the five subjects are presented in Table 3, as is the sum of these cross products. The sum divided by $n - 1$ (3.999947/4) equals, within rounding error, the actual R_C result reported in Table 2 for Function I.

An alternative presentation is graphic. Figure 1 presents the scattergram in which the five pairs of synthetic variable scores from Table 3 are arrayed. For example, note that the first subject's scores in Table 3 indicate that this subject is represented by the asterisk in the upper right position within the scattergram. Figure 1 also presents the least squares regression line best fitting these asterisks. In the two variable case, since the synthetic variables have means of zero, the slope of this regression line equals a beta weight, also equals the bivariate correlation between the synthetic variables, also equals the canonical correlation coefficient, i.e., .99999.

<u>INSERT FIGURE 1 ABOUT HERE.</u>

Table 4 presents computations that illustrate the meaning of two other canonical results, structure coefficients and index coefficients. Structure coefficients have the same meaning in a canonical analysis as in other analyses, i.e., structure coefficients are bivariate correlation coefficients between a given criterion or predictor variable and the synthetic variable involving the variable set to which the variable belongs. For example, since "ZX" was a criterion variable, the correlation

between "ZX" and "CRITCOMP" is the structure coefficient for "ZX." Note that the sum of the cross products of "ZX" and "CRITCOMP", labelled "XSTRUC" in Table 4, once divided by _n_ - 1, equals within rounding error the structure coefficient for "ZX" presented in Table 2. An index coefficient is the correlation coefficient between a variable and the synthetic variable consisting of variables from the variable set to which the variable does _not_ belong. Table 4 illustrates the calculation of the index coefficient for "ZX" on Function I. Thompson (1984) discusses the importance of these coefficients in greater detail.

INSERT TABLE 4 ABOUT HERE.

Three Common Interpretation Fallacies

The challenge of interpreting canonical results can give pause to even the most seasoned analyst. As Thompson (1980, pp. 1, 16-17) notes, one

> reason why the technique is rarely used involves the difficulties which can be encountered in trying to interpret canonical results... The neophyte student of canonical correlation analysis may be overwhelmed by the myriad coefficients which the procedure produces... [But] canonical correlation analysis produces results which can be theoretically rich, and if properly implemented the procedure can adequately capture some of the complex dynamics involved in educational reality.

However, the interpretation of canonical results can be facilitated if three common interpretation fallacies are avoided.

Interpretation of Function Coefficients

In an artificial world of forced-choices, the analyst would interpret structure coefficients while ignoring function coefficients. Structure coefficients are the most helpful coefficients to consult when interpreting canonical results, although many researchers do not interpret and some do not even report structure coefficients. Since structure coefficients inform the researcher of the correlation between each variable and the synthetic variables, these coefficients are what inform the researcher regarding the meaning of what is actually being correlated in a given analysis.

As noted previously, structure coefficients have the same meaning in the canonical cases as in the other analytic methods which the canonical methods subsume as special cases. For example, in principal components analysis the correlation between the scores on one variable and the factor scores on one factor is the structure coefficient for that variable on that factor. And as Gorsuch (1983, p. 207) notes, "the basic matrix for interpreting the factors is the factor structure." Similarly, in a discriminant analysis, the correlation between the scores on a predictor variable and the discriminant function scores on a given function is the structure coefficient for that variable on that function.

In the regression case, the correlation between scores on a predictor variable and the "YHAT" scores is the structure coefficient for the predictor variable. Just as structure coefficients are vitally important in interpreting results in

other analytic cases, structure coefficients can be very important in interpreting multiple regression results (Cooley & Lohnes, 1971, pp. 54-55). Thompson and Borrello (1985) present an explanation of this application and an actual research example in which the interpretation of beta weights rather than structure coefficients would conceivably have lead to incorrect conclusions.

Thus, with respect to canonical analysis, Meredith (1964, p. 55) suggested that, "If the variables within each set are moderately intercorrelated the possibility of interpreting the canonical variates by inspection of the appropriate regression weights [function coefficients] is practically nil." Similarly, Kerlinger and Pedhazur (1973, p. 344) argued that, "A canonical correlation analysis also yields weights, which, theoretically at least, are interpreted as regression [beta] weights. These weights [function coefficients] appear to be the weak link in the canonical correlation analysis chain." Levine (1977, p. 20, his emphasis) is even more emphatic:

> I specifically say that one *has* to do this [interpret structure coefficients] since I firmly believe as long as one wants information about the nature of the canonical correlation relationship, not merely the computation of the [synthetic function] scores, one must have the structure matrix.

The hypothetical results presented in Table 2 illustrate that the interpretation of only function coefficients can lead to

seriously distorted conclusions. The standardized function coefficients might lead the naive analyst to conclude that all four variables contribute appreciable information to the relationship between the two sets of synthetic variable scores on Function I. In reality, variables "ZY" and "ZB" share almost no variance at all with the function's scores.

Interpretation of Redundancy Coefficients

If the squared structure coefficients for a given set of variables are added and then the sum is divided by the number of variables in the set, the result informs the researcher regarding how much of the variance in the variables, on the average, is contained within the synthetic scores for that function. This result is called a variate adequacy coefficient (Thompson, 1984). Stewart and Love (1968) suggested that multiplying the adequacy coefficient times the squared canonical correlation yields a coefficient that they labelled a redundancy coefficient. Miller (1975) developed a partial test distribution to test the statistical significance of redundancy coefficients. Cooley and Lohnes (1976, p. 212) suggest that redundancy coefficients have great utility. In reality, the interpretation of redundancy coefficients does not make much sense in a conventional canonical analysis.

As Cramer and Nicewander (1979) proved in detail, redundancy coefficients are not truly multivariate. This is very disturbing, because the main argument in favor of multivariate methods (for both substantive and statistical reasons) is that these methods simultaneously consider all relationships during the analysis

(Thompson, 1986c).

Table 5 helps to illustrate the problem. The table presents the adequacy, redundancy, and squared Rc's for both functions for the hypothetical data, as well as the pooled values. For example, the pooled redundancy coefficient for the criterion variable set is .242783. Table 6 presents the results of four regression analyses for various criterion variables and predictor variable sets. The table illustrates that the average squared multiple R for a variable set equals the pooled redundancy coefficient for that variable set. The redundancy coefficient is the average of a set of univariate results!

A redundancy coefficient for a given variable set on a given function equals the adequacy coefficient for the set times the squared Rc for the function. The redundancy coefficient can only equal one when the synthetic variables for the function represent all the variance of every variable in the set, and the squared Rc also exactly equals one. This does not occur in practice, and it is difficult to conceive of any theoretical basis for ever formulating such an expectation. Thus, redundancy coefficients are useful only to test outcomes that rarely occur and which are not expected (Thompson, 1980, p. 16; Thompson, 1984). Furthermore, it seems contradictory to employ an analysis that use functions coefficients to optimize Rc, and then to interpret results not optimized as part of the analysis, i.e., redundancy coefficients.

Failure to Partition Using Canonical Commonality Analysis

Researchers have been aware for some time that

interpretation of regression results is often facilitated by conducting "commonality analyses" (Newton & Spurrell, 1967; Thompson, 1985a). These analyses partition variance to indicate how much variance is unique to a given variable, and how much variance is common to other variables. As an example analysis for the regression case, Seibold and McPhee (1979, pp. 364-365) present a cancer study in which the results would have been grossly misinterpreted if a commonality analysis had not been conducted.

Given that multiple regression is a special case of canonical correlation analysis, it seems reasonable to expect that the same variance partitioning procedures might also be very useful in the true canonical case. Although the details of the procedure are beyond the scope of the present paper, Thompson and Miller (1985) explain the procedure using an actual research example in which educators' perceptions of dying students and of death were investigated. It is suggested that the procedure may be very useful in research situations in which at least one of the variable sets consists of variables that are conceptually or theoretically distinct. As in the regression case, _the failure to employ commonality analysis can result in the distorted interpretation of results_.

Summary

In summary, the paper has briefly explained the logic underlying the basic calculations employed in canonical correlation analysis. Three common fallacious interpretation practices that may lead to incorrect conclusions based on

canonical results were presented. A small hypothetical data set was employed to make the discussion concrete.

Canonical correlation analysis is a powerful analytic method that frequently best honors the complex nature of the reality about which the researcher wishes to generalize. As Kerlinger (1973, p. 652) suggests, "some research problems almost demand canonical analysis." Similarly, Cooley and Lohnes (1971, p. 176) suggest that "it is the simplest model that can begin to do justice to this difficult problem of scientific generalization." However, the potentials of canonical correlation analysis will only be realized if researchers understand the logic underlying the method and if some serious interpretation pitfalls are avoided.

References

Campbell, D. T., & Erlebacher, A. (1975). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In M. Guttentag & E. L. Struening (Eds.), Handbook of evaluation research (Vol. 1). Beverly Hills: SAGE, pp. 597-617.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

Cooley, W. W., & Lohnes, P. R. (1971). Multivariate data analysis. New York: John Wiley and Sons.

Cooley, W. W., & Lohnes, P. R. (1976). Evaluation research in education. New York: Irvington.

Cramer, E. M., & Nicewander, W. A. (1979). Some symmetric, invariant measures of multivariate association. Psychometrika, 44, 43-54.

Goodwin, L. D., & Goodwin, W. L. (1985). Statistical techniques in AERJ articles, 1979-1983: The preparation of graduate students to read the educational research literature. Educational Researcher, 14, 5-11.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills: SAGE.

Harris, R. J. (April, 1987). A canonical cautionary. Paper presented at the annual meeting of the Society for Multivariate Experimental Psychology--Southwestern Division, New Orleans.

Jones, L. V., & Fiske, D. W. (1953). Models for testing

significance of combined results. Psychological Bulletin, 50, 375-381.

Kerlinger, F. N. (1973). Foundations of behavioral research (2nd ed.). New York: Holt, Rinehart and Winston.

Kerlinger, F. N., & Pedhazur, E. J. (1973). Multiple regression in behavioral research. New York: Holt, Rinehart and Winston.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. Psychological Bulletin, 85, 410-416.

Levine, M. S. (1977). Canonical analysis and factor comparison. Beverly Hills: SAGE.

Meredith, W. (1964). Canonical correlations with fallible data. Psychometrika, 29, 55-65.

Miller, J. K. (1975). The sampling distribution and a test for the significance of the bimultivariate redundancy statistic: A Monte Carlo Study. Multivariate Behavior Research, 10, 233-244.

Newton, R. G., & Spurrell, D. J. (1967). A development of multiple regression for the analysis of routine data. Applied Statistics, 16, 51-64.

Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hills: SAGE.

Seibold, D. R., & McPhee, R. D. (1979). Commonality analysis: A method for decomposing explained variance in multiple regression analyses. Human Communication Research, 5, 355-365.

Stewart, D. K., & Love, W. A. (1968). A general canonical

correlation index. Psychological Bulletin, 70, 160-163.

Thompson, B. (April, 1980). Canonical correlation: Recent extensions for modelling educational processes. Paper presented at the annual meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. ED 199 269)

Thompson, B. (November, 1981). The problem of OVAism. Paper presented at the annual meeting of the Mid-South Educational Research Association, Lexington, KY.

Thompson, B. (April, 1983). Teaching multivariate statistics as the bivariate case. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Beverly Hills: SAGE.

Thompson, B. (1985). Alternate methods for analyzing data from experiments. Journal of Experimental Education, 54, 50-55. (a)

Thompson, B. (April, 1985). Heuristics for teaching multivariate general linear model concepts. Paper presented at the annual meeting of the American Educational Research Association, Chicago. (ERIC Reproduction Service No. ED 262 073) (b)

Thompson, B. (1986). ANOVA versus regression analysis of ATI designs: An empirical investigation. Educational and Psychological Measurement, 46, 917-928. (a)

Thompson, B. (April, 1986). The place of qualitative research in contemporary social science: The importance of post-paradigmatic thought. Paper presented at the annual meeting of the American Educational Research Association, San

Francisco. (b)

Thompson, B. (November, 1986). Two reasons why multivariate methods are usually vital: An understandable reminder with concrete examples. Paper presented at the annual meeting of the Mid-South Educational Research Association, Memphis. (c)

Thompson, B. (April, 1987). The use (and misuse) of statistical significance testing: Some recommendations for improved editorial policy and practice. Paper presented at the annual meeting of the American Education Research Association, Washington, D.C.

Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

Thompson, B., & Miller, J. H. (January, 1985). A multivariate method of commonality analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin.

Table 1
Hypothetical "Bird Beak" Data

X	Y	A	B	ZX	ZY	ZA	ZB
9.0	7.0	10.0	8.0	-1.35287	-.63850	.00000	1.04326
11.0	6.0	10.0	5.0	-.44490	-1.26448	.00000	-.53744
12.0	8.0	8.0	4.0	.00908	-.01252	-.81650	-1.06434
13.0	10.0	8.0	5.0	.46306	1.23944	-.81650	-.53744
14.9	9.1	14.0	8.1	1.32563	.67606	1.63299	1.09595

Table 2
Selected Canonical Analysis Results

	Function I		Function II		
	Stn Fun	Struct	Stn Fun	Struct	Communality
ZX	-1.44986	-.69607	-.01281	.71798	1.000008
ZY	1.04101	-.00884	1.00924	.99996	.999998
ZA	-1.58021	-.61831	.02918	.78593	.999993
ZB	1.24215	.08146	.97723	.99983	1.006295
Rc	.99999		.02557		

Table 3
"Synthetic" Variate Scores for Function I

	ZX	ZY	ZA	ZB	CRITCOMP	PREDCOMP	CRITxPRED
	-1.35287	-.63850	.00000	1.04326	1.29678	1.29589	1.680484
	-.44490	-1.26448	.00000	-.53744	-.67129	-.66758	.448139
	.00908	-.01252	-.81650	-1.06434	-.02620	-.03183	.000833
	.46306	1.23944	-.81650	-.53744	.61889	.62266	.385358
	1.32563	.67606	1.63299	1.09595	-1.21819	-1.21913	1.485131
Sum							3.999947

Note. The sum of the cross-products (3.999947) divided by n-1 (4) is, within rounding error, the canonical correlation.

Table 4
Calculation of Structure and Index Coefficients

	ZX	CRITCOMP	PREDCOMP	XSTRUC	XINDEX
	-1.35287	1.29678	1.29589	-1.75438	-1.75317
	-.44490	-.67129	-.66758	.29866	.29701
	.00908	-.02620	-.03183	-.00024	-.00029
	.46306	.61889	.62266	.28658	.28833
	1.32563	-1.21819	-1.21913	-1.61487	-1.61612
Sum				-2.78425	-2.78424

Note. The sum of the cross-products of "ZX" and "CRITCOMP" (-2.78425) divided by n-1 (4) is (-.69606), within rounding error, the structure coefficient of "ZX" on Function I. The sum of the cross-products of "ZX" and "PREDCOMP" (-2.78424) divided by n-1 (4) is (-.69606), within rounding error, the index coefficient of "ZX" on Function I.

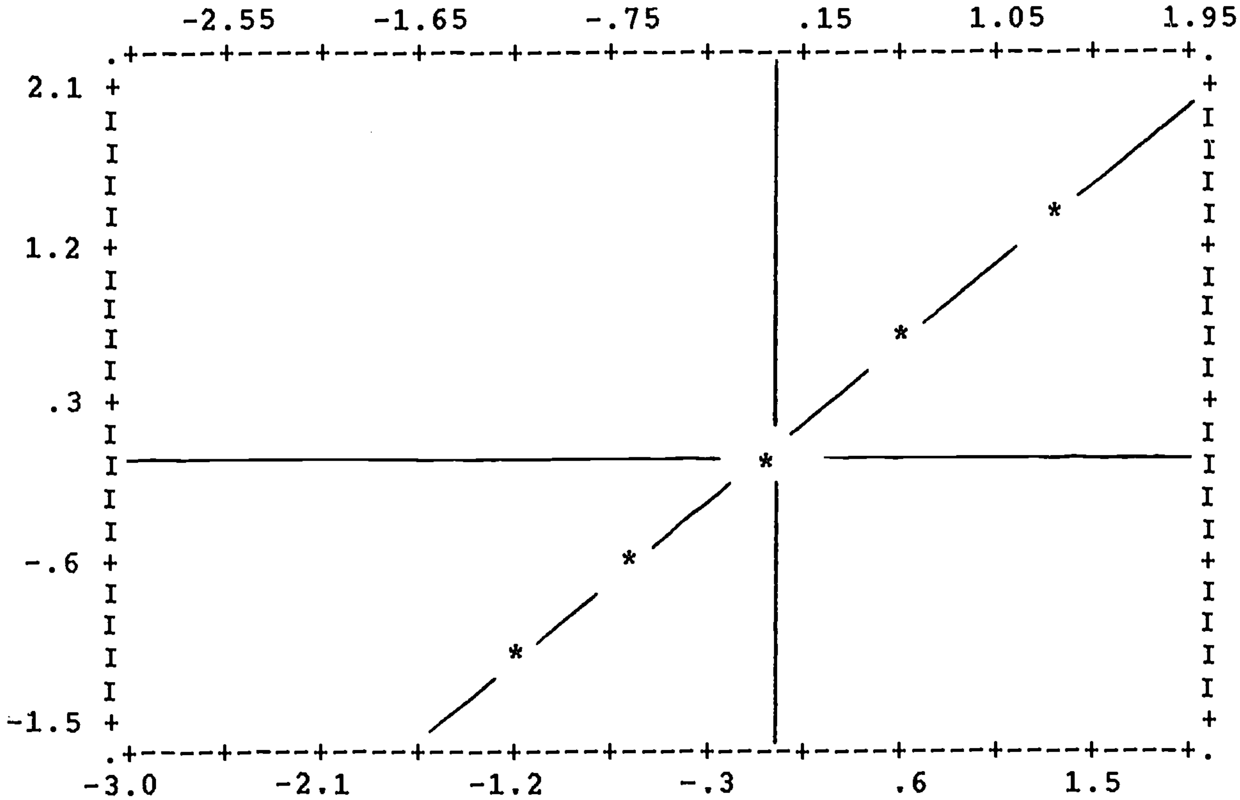
Table 5
Redundancy Calculations for Hypothetical Data

	Struc I	SQ	Struc II	SQ	Commun	Pooled Rd
	-0.69607	0.484513	0.71798	0.515495	1.000008	
	-0.00884	0.000078	0.99996	0.999920	0.999998	
SUM		0.484591		1.515415	2.000006	
Adequacy		0.242295		0.757707	1.000003	
Redundancy		0.242290		0.000492		0.242783
	-0.61831	0.382307	0.78593	0.617685	0.999993	
	0.08146	0.006635	0.99983	0.999660	1.006295	
SUM		0.388942		1.617345	2.006288	
Adequacy		0.194471		0.808672	1.003144	
Redundancy		0.194467		0.000525		0.194993
Rc SQ		0.99998		0.00065		

Table 6
Alternate Calculation of Pooled Coefficients

Criterion Variables	Predictor Variables	R	R SQ
X	A B	0.69630	0.48484
Y	A B	0.02705	0.00073
SUM			0.48557
Mean			0.24279
A	X Y	0.61864	0.38271
B	X Y	0.03153	0.00099
SUM			0.38370
Mean			0.19185

Figure 1
Scattergram of Canonical Composite Scores on Function I



APPENDIX A: Data and SPSS-X Command Files

File 'SWSMEP.DAT':

```
 90  70 100  80  5  6
110  60 100  50  0  3
120  80  80  40 10  5
130 100  80  50  0  7
149  91 140  81  5  4
```

File 'SWSMEP.SPS':

```
TITLE 'ANALYSIS OF adapted R.J. HARRIS DATA *****'
FILE HANDLE RJH/NAME='SWSMEP.DAT'
DATA LIST FILE=RJH/X 1-3 (1) Y 5-7 (1) A 9-11 (1) B 13-15 (1)
  A1 B1 16-21
COMPUTE ZX=(X-11.98)/2.20273
COMPUTE ZY=(Y-08.02)/1.59750
COMPUTE ZA=(A-10.00)/2.44949
COMPUTE ZB=(B-06.02)/1.89789
COMPUTE ZA1=(A1-4.0)/4.18330
COMPUTE ZB1=(B1-5.0)/1.58114
COMPUTE CRITCOMP=(-1.44986*ZX)+(1.04101*ZY)
COMPUTE PREDCOMP=(-1.58021*ZA)+(1.24215*ZB)
COMPUTE XSTRUC=ZX*CRITCOMP
COMPUTE XINDEX=ZX*PREDCOMP
PRINT FORMATS ZX TO XINDEX(F8.5)
LIST VARIABLES=ALL/CASES=50
CONDESCRIPTIVE X TO XINDEX
STATISTICS ALL
PEARSON CORR X TO XINDEX
REGRESSION VARIABLES=X TO B/CRITERIA=TOLERANCE(.000001)/
  DEPENDENT=X/ENTER A B
REGRESSION VARIABLES=X TO B/CRITERIA=TOLERANCE(.000001)/
  DEPENDENT=Y/ENTER A B
REGRESSION VARIABLES=X TO B/CRITERIA=TOLERANCE(.000001)/
  DEPENDENT=A/ENTER X Y
REGRESSION VARIABLES=X TO B/CRITERIA=TOLERANCE(.000001)/
  DEPENDENT=B/ENTER X Y
MANOVA X Y WITH A B/PRINT=CELLINFO(MEANS,COR)
  DISCRIM(RAW,STAN,COR,ALPHA(1.0))
  SIGNIF(DIMENR EIGEN MULTIV) ERROR(COR)/DESIGN/
SCATTERGRAM CRITCOMP (-6,3) WITH PREDCOMP (-3,6)
STATISTICS ALL
OPTIONS 4
MANOVA X Y WITH A1 B1/PRINT=CELLINFO(MEANS,COR)
  SIGNIF(DIMENR EIGEN MULTIV)
  DISCRIM(RAW,STAN,COR,ALPHA(1.00)) ERROR(COR)/DESIGN/
```