DOCUMENT RESUME

ED 428 081 TM 029 481

AUTHOR	Henson, Robin K.
TITLE	An Illustration That There Is a Multivariate Parametric General Linear Model: Canonical Correlation Analysis.
PUB DATE	1999-01-21
NOTE	38p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (San Antonio, TX, January 21-23, 1999).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Correlation; *Effect Size; *Heuristics; *Multivariate Analysis; Tables (Data)
IDENTIFIERS	*General Linear Model; Parametric Analysis; Statistical Package for the Social Sciences; Variables (Mathematics)

ABSTRACT

This paper illustrates how canonical correlation analysis can be employed to implement all the parametric tests that canonical methods subsume as special cases. The point is heuristic: all analyses are correlational, all apply weights to measured variables to create synthetic variables, and all yield effect sizes analogous to "r" squared. An appendix contains the Statistical Package for the Social Sciences command syntax for the analyses. (Contains 12 tables and 25 references.) (Author/SLD)

Running Head: CANONICAL AS THE GENERAL LINEAR MODEL

An Illustration that There is a Multivariate Parametric General Linear Model: Canonical Correlation Analysis

Robin K. Henson

Texas A&M University

Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, January 23, 1999.

Abstract

The paper illustrates how canonical correlation analysis can be employed to implement all the parametric tests that canonical methods subsume as special cases. The point is heuristic: *all* analyses are correlational, all apply weights to measured variables to create synthetic variables, and all yield effect sizes analogous to r^2.

An Illustration that There is a Multivariate Parametric General Linear Model: Canonical Correlation Analysis

In one of his innumerable seminal contributions, the late Jacob "Jack" Cohen (1968) demonstrated that multiple regression subsumes all the univariate parametric methods as special cases, and thus provides a univariate general linear model (GLM) that can be employed in all univariate analyses. At about the same time, researchers increasingly also came to realize that ANOVA was being overused, and in many cases used when other methods would have been more useful. One source of ANOVA overuse was that too many researchers erroneously associated ANOVA as an analysis with the ability to make causal statements when using experimental research designs; however, it is the design, and not the analysis that leads to the ability to make definitive causal statements!

As Humphreys (1978, p. 873, emphasis added) explained this phenomonon:

> The basic fact is that a measure of individual differences is not an independent variable [in an experimental design], and it does not become one by categorizing the scores and treating the categories as

> if they defined a variable under experimental control in a factorially designed analysis of variance.

Similarly, Humphreys and Fleishman (1974, p. 468) noted that categorizing variables in a nonexperimental design using an ANOVA analysis "not infrequently produces in both the investigator and his audience the illusion that he has experimental control over the independent variable. Nothing could be more wrong."

Furthermore, as Cliff (1987, p. 130, emphasis added) noted, the practice of discarding variance on intervally-scaled predictor variables in order to perform OVA analyses creates problems in almost all cases:

> Such divisions are not infallible; think of the persons near the borders. Some who should be highs are actually classified as lows, and vice versa. In addition, the "barely highs" are classified the same as the "very highs," even though they are different. Therefore, reducing a reliable variable to a dichotomy makes the variable <u>more unreliable</u>, not less.

These various realizations have led to less frequent use of OVA methods, and to more frequent use of general linear model approaches such as regression (cf. Edgington, 1974;

Elmore & Woehlke, 1988; Goodwin & Goodwin, 1985; Willson, 1980).

Since all analyses are correlational, and it is the design and not the analysis that yields the capacity to make causal inferences, the practice of converting intervally-scaled predictor variables to nominal scale so that ANOVA and other OVAs (i.e., ANCOVA, MANOVA, MANCOVA) can be conducted is inexcusable in many cases.

However, canonical correlation analysis, and not regression analysis, is the most general case of the general linear model (Baggaley, 1981, p. 129; Fornell, 1978, p. 168; Thompson, 1991, 1998). [Structural equation modeling (SEM) represents an even broader general linear model, but SEM is somewhat different in that this analysis usually also incorporates measurement error estimation as part of the analysis (cf. Bagozzi, Fornell, & Larcker, 1981; Fan, 1996, 1997).] In an important article, Knapp (1978, p. 410) demonstrated this in some detail and concluded that "virtually all of the commonly encountered tests of significance can be treated as special cases of canonical correlation analysis."

The present paper will illustrate how canonical correlation analysis can be employed to implement all the parametric tests that canonical methods subsume as special cases. The point is not that all research ought to be conducted with canonical analyses, rather the point is heuristic: all analyses are correlational, all analyses apply weights to measured variables to create synthetic variables that become the analytic focus, and all yield effect sizes analogous to r^2 that are important to interpret.

Understanding general linear model principles aids in realizing that parametric analyses are all fundamentally related. Individual methods, such as ANOVA or t-tests, can then be viewed from a global perspective which will, hopefully, facilitate thoughtful researcher judgment in selecting analyses as opposed to employing "lock-step" decision strategies that limit the utility of analyses.

The Basics of Canonical Correlation Analysis

The theory of canonical correlation analysis (CCA) has been with us for considerable time (Hotelling, 1935), but did not come into practical use until the onset of computerization (Krus, Reynolds, & Krus, 1976). This utility lies in being able to simultaneously examine all the relationships among all the variables.

In canonical analysis, the variables are considered to be members of two or more (in practice, almost always two) variable sets (e.g., pretest and posttest scores, aptitude and achievement scores) - otherwise we would analyze the data with factor analysis so as to consider simultaneously all the relationships, but without considering the existence of variable sets. Each set will include more than one variable, otherwise we generally would use a Pearson *r* or regression analysis. As will be shown later, these analyses are essentially the same thing anyway! While a comprehensive discussion of CCA is beyond a scope of the present paper, the reader is referred to Thompson (1991) for an accessible and user-friendly treatment of CCA.

A CCA will yield many useful statistics, the most recognized of which is the canonical correlation (R_c). The canonical correlation describes the relationship between two synthetic variables that have been modeled from their respective variable sets by applying weights to the measured variables. A canonical correlation will be produced for each function (i.e., for each set of standardized canonical function coefficients and respective measured variables). The number of functions, each of which will be perfectly uncorrelated with the others, equals the number of variables in the smaller of the variable sets. The canonical correlation can be squared to yield a variance-accounted-for effect size ($\underline{R}_c^2$), or the percentage of variance explainable in the criterion variable set predictable with knowledge of the variance in the predictor set.

One advantage of CCA, and other multivariate methods, lies in its simultaneous examination of the variables of interest, thus reducing risk of experimentwise Type I error (Fish, 1988; Henson, in press; Thompson, in press). A second, and perhaps often overlooked, advantage is the flexibility of the analysis in looking at various research problems. One example of this versatility can be found in a measurement study involving multivariate criterion-related score validity (Sexton, McLean, Boyd, Thompson, & McCormick, 1988). Thus, CCA can be used in either substantive or measurement inquiries.

Canonical Correlation Analysis as the General Linear Model

An heuristic data set for 12 elementary, middle, and high school students will be used to illustrate that CCA can conduct the other parametric methods that it subsumes, both univariate and multivariate alike. CCA will be used to perform a *t*-test, Pearson correlation, multiple regression, ANOVA, MANOVA, and descriptive discriminant analysis. Table 1 lists heuristic data on four intervally scaled variables related to motivational and personality issues: attributions of effort (EFFORT), attributions of ability (ABILIT), locus of control (LOCUS), and degree of extroversion (EXTROV). Also included are grouping data indicating some experimental treatment (TREAT) and whether students are in elementary, middle, or high school (GRADE). The reader will also notice five planned contrast variables which will be described later.

Analyses will be run using the SPSS (v6.1.4) statistics package. The command syntax for these analyses is included in Appendix A. Note that CCA is conducted using the MANOVA command (again, suggesting that these analyses must be related). Using Table 1 variable names, the SPSS commands for CCA are:

```
MANOVA
     LOCUS EXTROV WITH EFFORT ABILIT
     /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
     /DISCRIM=(STAN ESTIM COR ALPHA(.99)).
```

The SAS statistical software has a more direct command for CCA: PROC CANCORR. An example of SAS syntax used to perform a similar heuristic illustration can be found in Campbell and Taylor (1996).

INSERT TABLE 1 ABOUT HERE

Conducting t-test with Canonical

One of the most basic of statistical analyses is the *t*-test which is used to compare means between groups. Here a *t*-test was used to evaluate if the treatment and control groups (TREAT) differed on the EFFORT variable. Results reported in Table 2 indicate that the means of the groups were not statistically significantly different, $t = .310$, $p = .760$. A canonical analysis on the same variables yielded $F(1, 10) = .100$, $p = .760$. Table 2 also reports the CCA results, including the canonical correlation (R_c), squared canonical correlation (R_c^2), and Wilks lambda (λ). Wilks lambda, like R_c^2 is a variance-accounted-for type statistic. However, Wilks lambda indicates the variance not accounted for in the canonical correlation, modeled by $(1 - R_c^2)$. It is used for testing the statistical significance of R_c. As the magnitude of λ decreases (ranging from 0 to 1), the effect size (R_c^2) increases as does the likelihood of obtaining statistical significance.

Note that the *p* calculated values are identical between analyses. The test statistics (*t* and *F*) are different only in metric. In fact, the *F* distribution consists of squared values of the *t* distribution. Squaring $t = .310$ produces .096 which does match the *F* value. The slight difference in the values is arbitrary and solely due to rounding error by the statistics program.

INSERT TABLE 2 ABOUT HERE

Conducting Pearson Correlation with Canonical

When examining relationships between two variables, a Pearson correlation (*r*) is often invoked. The reader should immediately note conceptual similarities between a Pearson *r* and canonical analysis, even before examining the results from the SPSS analysis. Both investigate relationships between variables, only in the canonical case the measured variables of interest occur within multivariate sets.

A Pearson *r* was computed for EFFORT and ABILIT. Table 3 reports the obtained results, $r = -.6150$, $p < .05$. The CCA computed a squared canonical correlation coefficient of .378. A simple transformation of $R_c^2 = .378$ gives us $R_c = .6148$. The values are identical, save for rounding error and the fact that a canonical correlation cannot be negative. This is because the weights that are used in CCA scale the variables in the same direction, as such R_c will always range from 0 to 1. Note that the *p* values vary only because SPSS reports them differently; precise calculated *p* values will be identical.

Herein lies the most fundamental of general linear model principles: all analyses are correlational. The canonical correlation is nothing more than a bivariate *r* between the synthetic variables created in CCA after the application of weights. As Thompson (1991, p. 81) noted, "This conceptualization is appealing, because most researchers feel very comfortable thinking in terms of the familiar bivariate correlation coefficient."

Since the present heuristic CCA only had one variable in each set, the synthetic variables reflected the same relationship as did a Pearson *r* between the variables without the application of weights. This result should not be surprising, given the fact that multiplicative constants do not affect the value of *r*. The only effect the weights had in this case was to scale the variables in the same direction, thus yielding a positive value for R_c.

INSERT TABLE 3 ABOUT HERE

Conducting Multiple Regression with Canonical

As Cohen (1968) indicated, multiple regression subsumes all other univariate parametric analyses as special cases.

Therefore, there is a directly analogous relationship between Pearson *r* and multiple regression. Since CCA subsumes Pearson *r*, it should be apparent that it will do the same for multiple regression.

A multiple regression analysis was conducted with EFFORT being predicted by LOCUS and EXTROV. SPSS results of the regression and canonical analyses are found in Table 4. Again, all parallel statistics match within rounding error, with the exception of the weights. As with the *t* and *F* distribution above, the difference here is arbitrary. Beta (*B*) weights and standardized function coefficients are easily converted into each other using the following formulas:

$B \,/\, R_c$ = Function Coefficient

Function Coefficient * $R = B$

For example, LOCUS had a *B* weight of -.171156. Using R_c = .828 from the CCA, we find that the standardized function coefficient matches, within rounding error, that reported in Table 4 (-.171156 / .828 = -.2067). Since we know that the regression multiple *R* equals the canonical R_c, we can use the formulas above to find canonical function coefficients using only a regression analysis and *B* weights using only CCA!

INSERT TABLE 4 ABOUT HERE

Conducting Factorial ANOVA with Canonical

The SPSS syntax file (Appendix A) includes commands to compute the five orthogonal contrast variables reported in the Table 1 data. Planned contrasts can be used with OVA methods to test specific, theory-driven hypotheses as against omnibus hypotheses (Thompson, 1994). One advantage of using planned contrasts is the ease of pinpointing statistically significant effects without having to conduct post-hoc tests which include Bonferroni-type corrections for experimentwise error. It is important to note that the contrasts will yield the same overall effect [i.e., Sum of Squares (SOS) explained] as the omnibus test. They are necessary here to show that CCA can conduct ANOVA.

In the present analysis, a 3 X 2 factorial ANOVA was conducted with TREAT and GRADE as independent variables and EFFORT as the dependent variable. For the CCA, the contrast variables from Table 1 were used. The total number of contrasts that can be created equals the degrees of freedom for each main effect. The GRADE main effect has two degrees of freedom and is represented by CGR1 and CGR2. The TREAT main effect is represented by CTREAT with one degree of freedom. CTRGR1 and CTRGR2 are simply cross products of the other main effects and test the GRADE X TREAT interaction effects. Table 5 presents results for the ANOVA: GRADE, F = 19.367; TREAT, F = .510; GRADE X TREAT, F = 3.449. Note that the effect size (r^2) for the error term was .1323.

Obtaining comparable results with CCA requires us to conduct canonical analyses in four separate designs, using EFFORT as the dependent measure and the contrasts as independent variables. Design 1 included all planned contrasts, CGR1, CGR2, CTREAT, CTRGR1, and CTRGR2, to test the total effect (SOS explained). Design 2 used CTREAT, CTRGR1, and CTRGR2 to jointly test the TREAT and interaction effects. Design 3 used CGR1, CGR2, CTRGR1, and CTRGR2 to jointly test the GRADE and interaction effects. The final CCA, Design 4, used CGR1, CGR2, and CTREAT to jointly test the GRADE and TREAT effects. Table 6 displays the Wilks lambda values for each design. Remember that λ is something of a "reverse" effect size and will equal the effect for the error term. A quick comparison of λ for the total effect (Table 6) with the error effect size (Table 5) confirms this relationship between the statistics.

INSERT TABLE 6 ABOUT HERE

After canonical lambdas have been attained, we must use them to determine the omnibus ANOVA lambdas. This was done by dividing the Design 1 total effect (lambda) by the lambdas of the other designs. For example, to find the omnibus lambda for the GRADE main effect the total lambda (.11507) was divided by the Design 2 lambda (.85793), which reflects the joint effect of the contrast variables for the TREAT main effect and the GRADE X TREAT interaction effect. This process "removes" the effects of the other hypotheses, leaving the omnibus lambda for the GRADE main effect to be .13412516 (.11507 / .85793 = .13412516 = $\underline{\lambda}$). The same process was used to find the other ANOVA lambdas with results reported in Table 7.

INSERT TABLE 7 ABOUT HERE

One final step remained. ANOVA lambdas were converted into ANOVA $\underline{F}$ statistics using the following formula:

[(1 - Lambda) / Lambda] * ($\underline{df}$ error / $\underline{df}$ effect) = $\underline{F}$

To illustrate, the $\underline{F}$ value for the GRADE main effect was modeled by [(1 - .13412516) / .13412516] * (6 / 2) = 19.3671681. Table 8 reports transformations for both main effects and the interaction. Note that the *F* statistics obtained by the canonical process match those obtained by the factorial ANOVA (see Table 5), within rounding error of course.

INSERT TABLE 8 ABOUT HERE

Conducting Factorial MANOVA with Canonical

Since SPSS actually uses the MANOVA command to perform CCA, the two are obviously related. To illustrate the relationship, a 3 X 2 factorial MANOVA was computed with EFFORT and ABILIT as dependent variables and GRADE and TREAT as independent measures. Results from this analysis are found in Table 9. Since MANOVA is a multivariate method, Wilks lambdas are reported by SPSS and are used to test statistical significance.

The comparable canonical analysis was performed using the same process as with the ANOVA above. Four CCA designs using the contrast variables were run with canonical lambdas reported in Table 10. The subsequent conversion of these values to MANOVA lambdas is found in Table 11. The reader will note the equivalence of the MANOVA λs in Table 9 with those obtained through the canonical analysis in Table 11. The final conversion to *F* values wass not necessary here since the MANOVA uses the λ value to calculate F statistics, as against the SOS value in ANOVA.

INSERT TABLES 9 - 11 ABOUT HERE

Conducting Discriminant Analysis with Canonical

Discriminant analysis is a multivariate method that can either be used predictively to classify persons into groups or descriptively where variables identify latent structures among groups (Huberty, 1994). The descriptive discriminant analysis (DDA) case is especially useful as the preferred substitute for a one-way MANOVA or as a post hoc analysis to multi-way MANOVA analyses.

To demonstrate the DDA and CCA relationship, a descriptive discriminant analysis was conducted with TREAT as the nominally scaled predictor variable and EFFORT and ABILIT as criterion variables. Table 12 reports a non-statistically significant result $\chi^2(2, 9) = .648$, $p = .723$. The canonical analysis was conducted using the planned contrast variable CTREAT as the predictor. Results of the CCA are also reported in Table 12. The reader will note that the analyses yield identical results. One arbitrary difference is in the reporting of a χ^2 statistic for the discriminant analysis as opposed to the CCA *F* value. As with the *t* and *F* distributions described above, the difference is arbitrary since the χ^2 and *F* statistics represent the same value expressed in a different metric.

INSERT TABLE 12 ABOUT HERE

Conclusion

The purpose of the present paper has been to illustrate that canonical correlation analysis represents the multivariate parametric general linear model. As such, CCA can be used to conduct the univariate and multivariate analyses that CCA subsumes. The point is heuristic and not intended to suggest that all analyses should be conducted with CCA. In fact, it is quite clear in the ANOVA and MANOVA examples that CCA, at least as reported by SPSS, is the long way to the same results. However, CCA would be superior to ANOVA and MANOVA when the independent variables are intervally scaled, thus eliminating the need to discard variance.

Knowing that there is a general linear model and understanding that all parametric analyses are intricately related can be of great educational value to both students and teachers of quantitative methods. Knowing these relationships facilitates understanding of commonalities and differences among all the parametric methods.

References

Baggaley, A. R. (1981). Multivariate analysis: An introduction for consumers of behavioral research. Evaluation Review, 5, 123-131.

Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of a structural relations model. Multivariate Behavioral Research, 16, 437-454.

Campbell, K. T., & Taylor, D. L. (1996). Canonical correlation analysis as a general linear model: A heuristic lesson for teachers and students. Journal of Experimental Education, 64, 157-171.

Cliff, N. (1987). Analyzing multivariate data. San Diego, Harcourt Brace Jovanovich.

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

Edgington, E. S. (1974). A new tabulation of statistical procedures used in APA journals. American Psychologist, 29, 25-26.

Elmore, P. B., & Woehlke, P. L. (1988). Statistical methods employed in American Educational Research Journal, Educational Researcher, and Review of Educational Research from 1978 to 1987. Educational Researcher, 17(9), 19-20.

Fan, X. (1996). Canonical correlation analysis as a general analytic model. In B. Thompson (Ed.), Advances in social science methodology (Vol. 4, pp. 71-94). Greenwich, CT: JAI Press.

Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common? Structural Equation Modeling, 4, 65-79.

Fish, L. J. (1988). Why multivariate methods are usually vital. Measurement and Evaluation in Counseling and Development, 21, 130-137.

Fornell, C. (1978). Three approaches to canonical analysis. Journal of the Market Research Society, 20, 166-181.

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(2), 5-11.

Henson, R. K. (in press). Multivariate normality: What is it and how is it assessed? In B. Thompson (Ed.), Advances in social science methodology (Vol. 5). Greenwich, CT: JAI Press.

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

Huberty, C. (1994). Applied discriminant analysis. New York: Wiley.

Humphreys, L. G. (1978). Doing research the hard way: Substituting analysis of variance for a problem in correlational analysis. Journal of Educational Psychology, 70, 873-876.

Humphreys, L. G., & Fleishman, A. (1974). Pseudo-orthogonal and other analysis of variance designs involving individual-differences variables. Journal of Educational Psychology, 66, 464-472.

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

Krus, D. J., Reynolds, T. S., & Krus, P. H. (1976). Rotation in canonical variate analysis. Educational and Psychological Measurement, 36, 725-730

Sexton, J. D., McLean, M., Boyd, R. D., Thompson, B., & McCormick, K. (1988). Criterion-related validity of a new standardized measure for use with infants who are handicapped. Measurement and Evaluation in Counseling and Development, 21, 16-24.

Thompson, B. (in press). Canonical correlation analysis. In L. Grimm & P. Yarnold (Eds.), Reading and understanding multivariate statistics (Vol. 2). Washington, DC: American Psychological Association.

Thompson, B. (1991). A primer on the logic and use on canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24(2), 80-95.

Thompson, B. (1994). Planned versus unplanned and orthogonal versus nonorthongonal contrasts: The neo-classical perspective. In B. Thompson (Ed.), Advances in social science methodology (Vol. 3, pp. 3-27). Greenwich, CT: JAI Press.

Thompson (1998, April). Five methodology errors in educational research: The pantheon of statistical significance and other faux pas. Invited address presented at the annual meeting of the American Educational Research Association, San Diego. (Internet URL http://acs.tamu.edu/~bbt6147/)

Willson, V. L. (1980). Research techniques in AERJ articles: 1969 to 1978. Educational Researcher, 9(6), 5-10.

Table 1

Heuristic Data (n=12) for Canonical Correlation Illustration.

ID	EFFORT	ABILIT	LOCUS	EXTROV	GRADE	TREAT	CGR1	CGR2	CTREAT	CTRGR1	CTRGR2
1	10	12	18	15	1	1	-1	-1	-1	1	1
2	15	14	19	16	1	1	-1	-1	-1	1	1
3	17	18	18	13	1	2	-1	-1	1	-1	-1
4	14	13	15	10	1	2	-1	-1	1	-1	-1
5	09	15	14	04	2	1	0	2	-1	0	-2
6	06	19	16	04	2	1	0	2	-1	0	-2
7	06	20	12	07	2	2	0	2	1	0	2
8	07	19	16	03	2	2	0	2	1	0	2
9	18	11	06	18	3	1	1	-1	-1	-1	1
10	17	10	04	13	3	1	1	-1	-1	-1	1
11	12	09	10	12	3	2	1	-1	1	1	-1
12	14	13	09	14	3	2	1	-1	1	1	-1

Table 2

Conducting *t*-test with Canonical (EFFORT by TREAT).

t-test Analysis		Canonical Analysis	
$t(10)$	.31	$F(1,10)$	.10
p	.076	p	.076
M (TREAT 1)	12.500		
SD	4.848	R_c	.100
M (TREAT 2)	11.6667	R_c^2	.010
SD	4.320	lambda	.990

Table 3

Conducting Pearson Correlation with Canonical (EFFORT by ABILITY).

Pearson *r* Analysis		Canonical Analysis	
r	-.6150	R_c	.6148
		R_c^2	.378
		lambda	.622
$p <$	.05	p	.033

Note. R_c cannot be negative. Calculated p values vary only due to reporting style of statistical program.

Table 4

Conducting Multiple Regression with Canonical (EFFORT by LOCUS and ABILIT).

Multiple Regression Analysis		Canonical Analysis	
$\underline{R}$	.8278	$\underline{R}_c$	.828
$\underline{R}^2$	.68525	$\underline{R}_c^2$	.685
		lambda	.31475
$\underline{F}(2, 9)$	9.79712	$\underline{F}(2, 9)$	9.79712
$\underline{p}$	.0055	$\underline{p}$	.006
Beta Weights		Function Coefficients	
LOCUS	-.171156	LOCUS	-.207
EXTROV	.766543	EXTROV	.926

Table 5

3 X 2 Factorial ANOVA (EFFORT by GRADE and TREAT).

Source	SOS	$\underline{df}$	MS	$\underline{F}$	$\underline{p}$	$\underline{r}^2$
GRADE	158.167	2	79.083	19.367	.002	74.29%
TREAT	2.083	1	2.083	.510	.502	.98%
G X T	28.167	2	14.083	3.449	.101	13.23%
Error	24.500	6	4.083			
Total	212.917	11				

Table 6

Canonical Analyses on Four Designs (EFFORT by Contrasts).

Design	Independent Variables	lambda
1	CGR1, CGR2, CTREAT, CTRGR1, CTRGR2	.11507
2	CTREAT, CTRGR1, CTRGR2	.85793
3	CGR1, CGR2, CTRGR1, CTRGR2	.12485
4	CGR1, CGR2, CTREAT	.24736

Table 7

Conversion of Canonical Lambdas to Omnibus ANOVA Lambdas.

ANOVA Effect	Designs	Transformation	ANOVA lambda
GRADE	1 / 2	.11507/.85793	.13412516
TREAT	1 / 3	.11507/.12485	.92166600
GRADE X TREAT	1 / 4	.11507/.24736	.46519243

Table 8

Conversion of ANOVA Lambdas ANOVA F Statistics.

Source	Transformation	F
GRADE	[(1 - .13412516) / .13412516] * (6 / 2) =	19.3671681
TREAT	[(1 - .92166600) / .92166600] * (6 / 1) =	.5099504
GRADE X TREAT	[(1 - .46519243) / .46519243] * (6 / 2) =	3.4489441

Table 9

3 X 2 Factorial MANOVA (EFFORT and ABILIT by GRADE and TREAT).

Source	lambda	*df*	*F*	*p*
GRADE	.05061	4, 10	8.61299	.003
TREAT	.61798	2, 5	1.54541	.300
GRADE X TREAT	.44653	4, 10	1.24122	.354

Table 10

Canonical Analyses on Four Designs (EFFORT and ABILIT by Contrasts).

Design	Independent Variables	lambda
1	CGR1, CGR2, CTREAT, CTRGR1, CTRGR2	.03184
2	CTREAT, CTRGR1, CTRGR2	.62924
3	CGR1, CGR2, CTRGR1, CTRGR2	.05153
4	CGR1, CGR2, CTREAT	.07132

Table 11

Conversion of Canonical Lambdas to Omnibus MANOVA Lambdas.

MANOVA Effect	Designs	Transformation	MANOVA lambda
GRADE	1 / 2	.03184/.62924	.05060072
TREAT	1 / 3	.03184/.05153	.61789249
GRADE X TREAT	1 / 4	.03184/.07132	.44643859

Table 12

Conducting Discriminant Analysis with Canonical (EFFORT and ABILIT by TREAT).

Discriminant Analysis		Canonical Analysis	
R_c	.2636	R_c	.264
R_c^2	.0695	R_c^2	.069
lambda	.930518	lambda	.93052
χ^2	.648	F	.33602
df	2, 9	df	2, 9
p	.7232	p	.723

Appendix A

```
TITLE ' Illustrating that there is a parametric general '.
TITLE ' linear model: Canonical correlation analysis '.
TITLE ' Robin K. Henson - SERA1999 '.
COMMENT **************************************************
COMMENT Heuristic data for 12 cases
COMMENT EFFORT - attributions of effort
COMMENT ABILIT - attributions of ability
COMMENT LOCUS - external vs internal locus of control
COMMENT EXTROV - degree of extroversion scale
COMMENT GRADE - elementary(1), middle(2), high(3) school
COMMENT TREAT - treat(1), control(2) groups.
SET BLANKS=SYSMIS UNDEFINED=WARN PRINTBACK LISTING.
DATA LIST
     FILE='c:\presentations\ccaasglm_data_sera99.txt' FIXED RECORDS=1
     /ID 1-2 EFFORT 4-5 ABILIT 7-8 LOCUS 10-11 EXTROV 13-14 GRADE 16
     TREAT 18.
EXECUTE.
COMMENT Show that cca can do t-test.
T-TEST
  GROUPS=TREAT(1 2)
  /MISSING=ANALYSIS
  /VARIABLES=EFFORT
  /CRITERIA=CIN(.95) .
MANOVA
   TREAT WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
COMMENT Show that cca can do Pearson r.
CORRELATIONS
  /VARIABLES=EFFORT ABILIT
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE .
MANOVA
   EFFORT WITH ABILIT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
COMMENT Show that cca can do multiple regression.
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT EFFORT
  /METHOD=ENTER LOCUS EXTROV  .
MANOVA
   LOCUS EXTROV WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
COMMENT Show that cca can do factorial ANOVA.
COMMENT Compute contrast varables to do cca.
IF (GRADE = 1) CGR1 = -1.
IF (GRADE = 2) CGR1 = 0.
IF (GRADE = 3) CGR1 = 1.
COMMENT Tests equality of the means of elementary(4) vs high school(4) students.
EXECUTE.
```

```
IF (CGR1 = -1) CGR2 = -1.
IF (CGR1 = 0) CGR2 = 2.
IF (CGR1 = 1) CGR2 = -1.
EXECUTE.
COMMENT Tests equality of means of middle(4) vs elementary high school(8)
students.
IF (TREAT = 1) CTREAT = -1.
IF (TREAT = 2) CTREAT = 1.
EXECUTE.
COMMENT Tests equality of means of treatment (6) vs control groups (6).
COMPUTE CTRGR1 = CGR1 * CTREAT.
COMPUTE CTRGR2 = CGR2 * CTREAT.
EXECUTE.
COMMENT Tests treatment by grade interaction effects.
COMMENT Show contrast variables are orthogonal.
CORRELATIONS
  /VARIABLES=CGR1 CGR2 CTREAT CTRGR1 CTRGR2
  /PRINT=TWOTAIL SIG
  /MISSING=PAIRWISE .
COMMENT Step one: run factorial ANOVA and cca on constrast variables.
ANOVA
  VARIABLES=EFFORT
  BY GRADE(1 3) TREAT(1 2)
  /MAXORDERS ALL
  /METHOD UNIQUE
  /FORMAT LABELS .
MANOVA
   CGR1 CGR2 CTREAT CTRGR1 CTRGR2 WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CTREAT CTRGR1 CTRGR2 WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CGR1 CGR2 CTRGR1 CTRGR2 WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CGR1 CGR2 CTREAT WITH EFFORT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
COMMENT Show cca can do MANOVA.
MANOVA
  EFFORT ABILIT BY GRADE(1 3) TREAT(1 2)
  /PRINT SIGNIF(MULT UNIV )
  /NOPRINT PARAM(ESTIM)
  /METHOD=UNIQUE
  /ERROR WITHIN+RESIDUAL
  /DESIGN  .
MANOVA
   CGR1 CGR2 CTREAT CTRGR1 CTRGR2 WITH EFFORT ABILIT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CTREAT CTRGR1 CTRGR2 WITH EFFORT ABILIT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
```

```
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CGR1 CGR2 CTRGR1 CTRGR2 WITH EFFORT ABILIT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
MANOVA
   CGR1 CGR2 CTREAT WITH EFFORT ABILIT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
COMMENT Show cca can do discriminant analysis.
DISCRIMINANT
  /GROUPS=TREAT(1 2)
  /VARIABLES=EFFORT ABILIT
  /ANALYSIS ALL
  /PRIORS  EQUAL
  /CLASSIFY=NONMISSING POOLED .
MANOVA
   EFFORT ABILIT WITH CTREAT
   /PRINT=SIGNIF (MULTIV EIGEN DIMENR)
   /DISCRIM=(STAN ESTIM COR).
```

TM029481

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: AN ILLUSTRATION THAT THERE IS A MULTIVARIATE PARAMETRIC GENERAL LINEAR MODEL: CANONICAL CORRELATION ANALYSIS	
Author(s): ROBIN K. HENSON	
Corporate Source:	Publication Date: 1/99

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here

Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ROBIN K. HENSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

or here

Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: X Robin K. Henson	Position: RES ASSOCIATE
Printed Name: Robin K. Henson	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1831
	Date: 1/13/99

OVER