Commonality Analysis: Understanding Variance Contributions to Overall Canonical Correlation Effects of Attitude Toward Mathematics on Geometry Achievement

Robert M. Capraro **Mary Margaret Capraro**
Texas A & M University

Canonical correlation analysis is the most general linear model subsuming all other univariate and multivariate cases (Kerlinger & Pedhazur, 1973; Thompson, 1985, 1991). Because “reality” is a complex place, a multivariate analysis such as canonical correlation analysis is demanded to match the research design. It is the purpose of this paper to increase the awareness and use of canonical correlation analysis and, specifically to demonstrate the value of the related procedure of commonality analysis. Commonality analysis provides the researcher with information regarding the variance explained by each of the measured variables and the common contribution from one or more of the other variables in a canonical analysis (Beaton, 1973; Frederick, 1999). This paper identifies confidence as contributing the most unique variance to the model, being more important than either intrinsic value or worry to geometry content knowledge and spatial visualization.

In developing the concept of commonality analysis (CA) one must be familiar with canonical correlation analysis (CCA), a multivariate technique. Most educational research settings demand an analysis that accounts for reality so a multivariate analysis should be used to match the research design as closely as possible. Canonical correlation analysis (CCA) is the most general case of the general liner model (GLM) (Baggaley, 1981). All univariate and multivariate cases can be treated as special cases of CCA (Thompson, 1984, 1991). As Henson (2000) noted, “CCA is superior to ANOVA and MANOVA when the independent variables are intervally scaled, thus eliminating the need to discard variance” otherwise one should refrain from using canonical correlation for these purposes.

There are several rational reasons for selecting CCA. Regarding OVA methods, the first is that CCA honors the relationship among variables because CCA does not require the variables to be converted from their original scale into arbitrary predictor categories (Frederick, 1999). Second, the method honors the reality to which the researcher is often trying to generalize (Henson, 2000; Tatsuoka, 1971; Thompson, 1984,1991). Third, reality has multiple outcomes with multiple causes; thus, it follows that most causes have multiple effects necessitating a multivariate approach (Thompson, 1991). Therefore, any analytic model that does not account for reality in which research is conducted distorts interpretations and potentially provides unreliable results (Tatsuoka, 1971). Historicalyr, research studies rarely used CCA. Prohibitive calculations, difficulty in trying to interpret canonical results and general unfamiliarity with the method contributed to CCA's absence from the literature (Baggaley, 1981; DeVito, 1976; Fan, 1996; Thompson, 1984).

Using CCA in real-life research situations increases the reliability of the results by limiting the inflation of Type I "experimentwise" error rates by reducing the number of analyses in a given study (Shavelson, 1988; Thompson, 1991). As Thompson (1991) stated CCA's limitation of "experimentwise" error, reduces the probability of making a Type I error anywhere within the investigation. Commonly, Type I error refers to "testwise" error rates, the probability of making an error in regards to a specified hypothesis test.

Thompson (1984) stated that some research almost demands CCA in which “ . . . it is the simplest model that can do justice to the difficult problem of scientific generalization” (p. 8). Furthermore, the use of CCA leads to the use of commonality analysis (Thompson, 1984). Although the voluminous output from CCA can be difficult to interpret (Tatsuoka, 1971; Thompson, 1984, 1990), however, once complete and noteworthy results emerge, one is obliged to consider the use of commonality analysis.

Commonality Analysis

Commonality analysis, also known as elements analysis and components analysis was developed for multiple regression analysis in the late 1960's (Newton & Spurell, 1967; Thompson, Miller, & James, 1985). Commonality analysis provides the researcher with information regarding the variance explained by each of the measured variables and the common contribution from one or more of the other variables (Beaton, 1973; Frederick, 1999). Partitioning of the variables takes two distinct forms. The first is in the form of explanatory ability that is in common with other variable(s). The second explanatory power can be attributed to unique contributions of a variable. This information should not be confused with interaction effects of regression. Interaction effects cannot be considered as indicating a unique contribution to the criterion set. Each variable in the predictor set simply adds predictive ability or increased variance to the first one (variable) entered. Commonality analysis, however, determines the variance explained that two or more predictor variables share that is useful in predicting relationships with the criterion variable set. Essentially, Beaton (1973) stated that CA partitions the common and unique variance of the several possible predictor variables on the set of criterion variables.

Commonalities can be either positive or negative. Beaton (1973) explained that negative commonalities are rare in educational research but more common in physical science research. While both positive and negative commonalities are useful, negative commonalities indicate that one variable confounds the variance explained by another. When referring to the power of CA, power is synonymous with variance explained. Negative commonalities may actually indicate improved power when both variables are used to make predictions (Beaton, 1973). The following example illustrates the relationship: An Olympic track athlete must be fast and strong, therefore, a strong-fast athlete would be correlated with success at running track. However, one would believe the two variables (fast and strong) would be moderately negatively correlated, that is as muscle strength and mass increase, speed would decrease. The negative commonality between speed and strength would indicate a confounded variable. In this case, by knowing both the speed and strength one would expect to make better predictions of successful track running. Imagine just knowing the speed or strength of the athlete. A fast athlete may perform well in a short sprint but be severely impaired in a distance event. Conversely, a strong athlete may excel in endurance and persevere for distance, but lack the speed to win. The negative commonality in this case indicates that the power of both variables is greater when the other variable is also used.

Conducting a Commonality Analysis

The complexity of conducting a CA ranges from the unsophisticated to the sublime. Frederick (1999) suggested the use of no more than four (predictor) measured variables because as the number of predictors increases so does the difficulty of interpretation. Frederick continued, explaining that the commonality calculations increase in difficulty exponentially as the number of predictors increases. Pedhazur (1982) and Frederick (1999) recommend that to avoid some of the complexities one should group similar variables or do some preliminary analyses to distinguish the most powerful predictors before conducting the CA such as a canonical correlation analysis.

The full model CCA is run with the following SPSS syntax:

```
MANOVA
spacerel gcksum with int.val worry confid
/print=signif (multiv eigen dimenr)
/discrim=(stan estim cor alpha(.999))/design.
```

The criterion variables are space relations (spacerel) and geometry content knowledge (gcksum). The predictor variables are confidence solving mathematics problems (confid), worry (worry), and finally mathematics intrinsic value (int.val). Possible relationships among variables are illustrated by Figure 1.

The Venn diagram illustrating commonality analysis in Figure 1 serves as a model for the comparison of data examined in the present paper. The data was collected in a southeastern state and represents 287 sixth grade students' scores on three measures, the *Space Relations* portion of the *Differential Aptitude Test* (Bennett, Seashore, & Wesman, 1973), the *Geometry Content Knowledge* test (Carroll, 1998), and the *Mathematics Attitude Scale* (Gierl & Bisanz, 1997).

The first step in running a CA begins with the findings of the CCA (the syntax provided earlier; also see the Appendix for the complete SPSS syntax). The next step involves running a descriptive analysis for the purposes of obtaining the standard deviation and means of each variable in order to calculate *z*-scores. The *z*-scores are computed for the observed variables by the following SPSS syntax:

COMPUTE zspace = (spacerel- mean)/standard deviation.
COMPUTE zgck = (gcksum- mean)/standard deviation.

To create the synthetic canonical variate scores, multiply the z-scores by the standardized canonical function coefficients (found in the original CCA), and then sum the scores for the function. The following SPSS syntax will yield the two sets of criterion variable composite scores (called crit1 and crit2) for both canonical functions.

COMPUTE crit1=(standardized canonical function coefficient I*zspace)
+(standardized canonical function coefficient I*zgck).
COMPUTE crit2=(standardized canonical function coefficient II*zspace)
+(standardized canonical function coefficient II*zgck).

Next, the CA requires running several multiple regression analyses for each criterion composite i.e., crit1 and crit2 using all possible combinations of predictor variables. Refer to Table 1 for the combinations for 2 or 3 predictor variables.

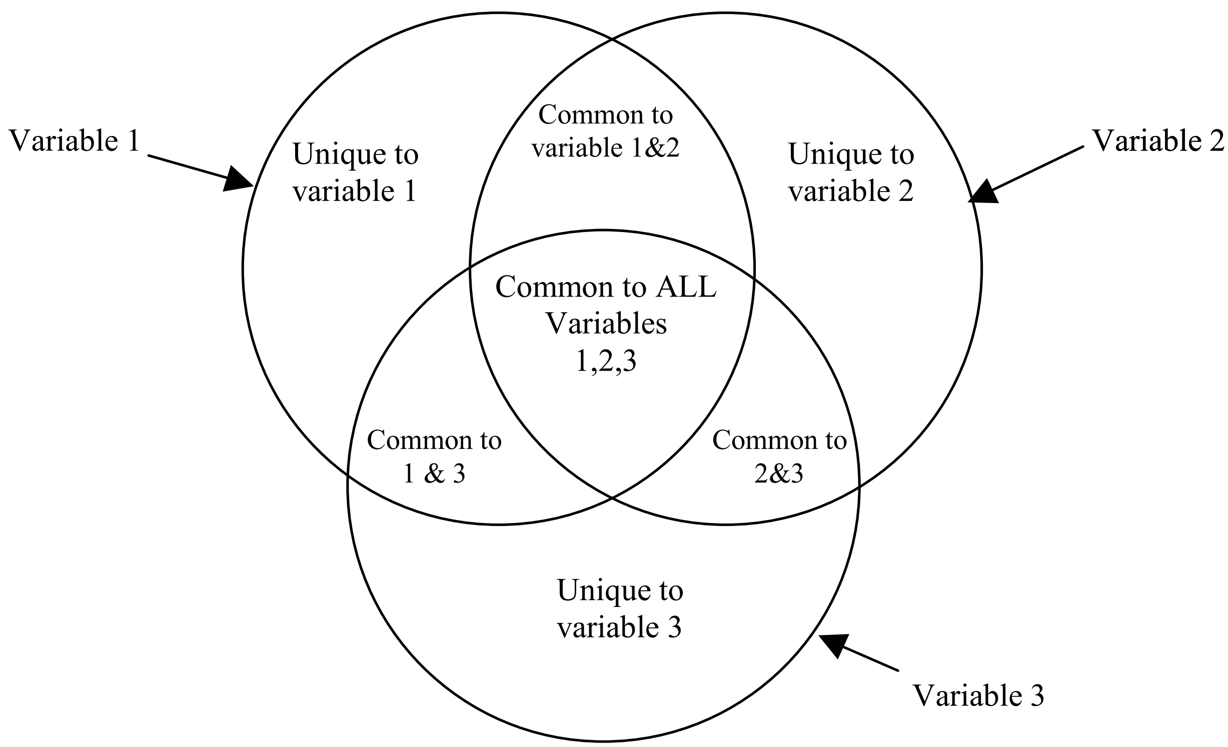

Figure 1. *Illustrating Commonality Analysis.*

Table 1. *Methods of Computing Unique and Common Variance.*

Two Predictor Variables
$U(1)=R^2_{12}-R^2_{2}$, $U(2)=R^2_{12}-R^2_{1}$, $C(12)=R^2_{2}+R^2_{1}-R^2_{12}$
Three Predictor Variables
$U(1)=R^2_{123}-R^2_{23}$, $U(2)=R^2_{123}-R^2_{13}$, $U(3)=R^2_{123}-R^2_{12}$, $C(12)=R^2_{13}-R^2_{3}+R^2_{23}-R^2_{123}$, $C(13)=R^2_{12}-R^2_{2}+R^2_{23}-R^2_{123}$, $C(23)=R^2_{12}-R^2_{1}+R^2_{13}-R^2_{123}$, $C(123)=R^2_{12}-R^2_{2}+R^2_{3}-R^2_{12}-R^2_{13}-R^2_{23}+R^2_{123}$

Note: U= unique variance, C= common variance, C13 = Common to variables 1 & 3
R^2=squared multiple correlation from the respective regression analysis.

Table 2. *Commonality Table.*

Variance	Function I				Function II			
Partition	Intrinsic	Worry	Confidence	Composite	Intrinsic	Worry	Confidence	Composite
U Intrinsic	0.001	0.001			0.019			0.019
U Worry	0.009		0.009		0	0		
U Confidence	0.188			0.188	.003		.003	
C IW	0.002	0.002			0.002			0.002
C IC	0.049	0.049	0.049	0.049	-0.003			-0.003
C WC	0.004		0.004	0.004	0			
C IWC	-0.011	-0.011	-0.011	-0.011	0	0	0	0
R^2 with Crit	0.041	0.004	0.230	0.242	0.018	0.002	0	0.021

Table 3. *Comparisons of Multivariate CCA and Univariate Multiple Regression with All Predictors.*

	Function	
Statistic	I	II
Multiple Regression (R^2)	.242	.021
Canonical Correlation (Rc^2)	.242	.021

Finally, add or subtract relevant effects to calculate the unique and common variance components for each predictor variable on each composite. Do this either by hand or by spreadsheet. The number of components in an analysis will equal (2^k-1), where k= number of predictor variables in the set. So, four predictors produce, 15 components, four-first order (unique), six-second order (common to two variables), four-third order (common to three variables), and one- fourth order (common to all).

The analysis of the present data consisted of two criterion variables, space relations and geometry content knowledge, and three predictor variables from the subscales of the *Mathematics Attitude Scales,* confidence, worry, and intrinsic value. One would expect, that through the application of (2^k-1), to have seven composites, three-first order (unique), three-second order (common to two) and one-third order (common to all). Results are displayed in Table 2.

Recall that both a full CCA and multiple linear regression with all predictors were conducted. The results displayed in Table 3 confirm that both procedures yielded the same results. Note that the R^2 and Rc^2 for Functions I and II are the same for both the multiple regression and CCA. The R^2 from the multiple linear regression reflect the additive effects of all the predictor combinations. These numbers will be confirmed again when summing all of the separate composites for each function (Table 2).

Analyzing Results

One must return to the Venn diagram (Figure 1) and then reconstruct it using the actual data from Table 2. This graphic helps one to visualize the relationships of the partitioned variance. If one only requires the variance explained from the entire CCA then there is no need to conduct a CA. However, the

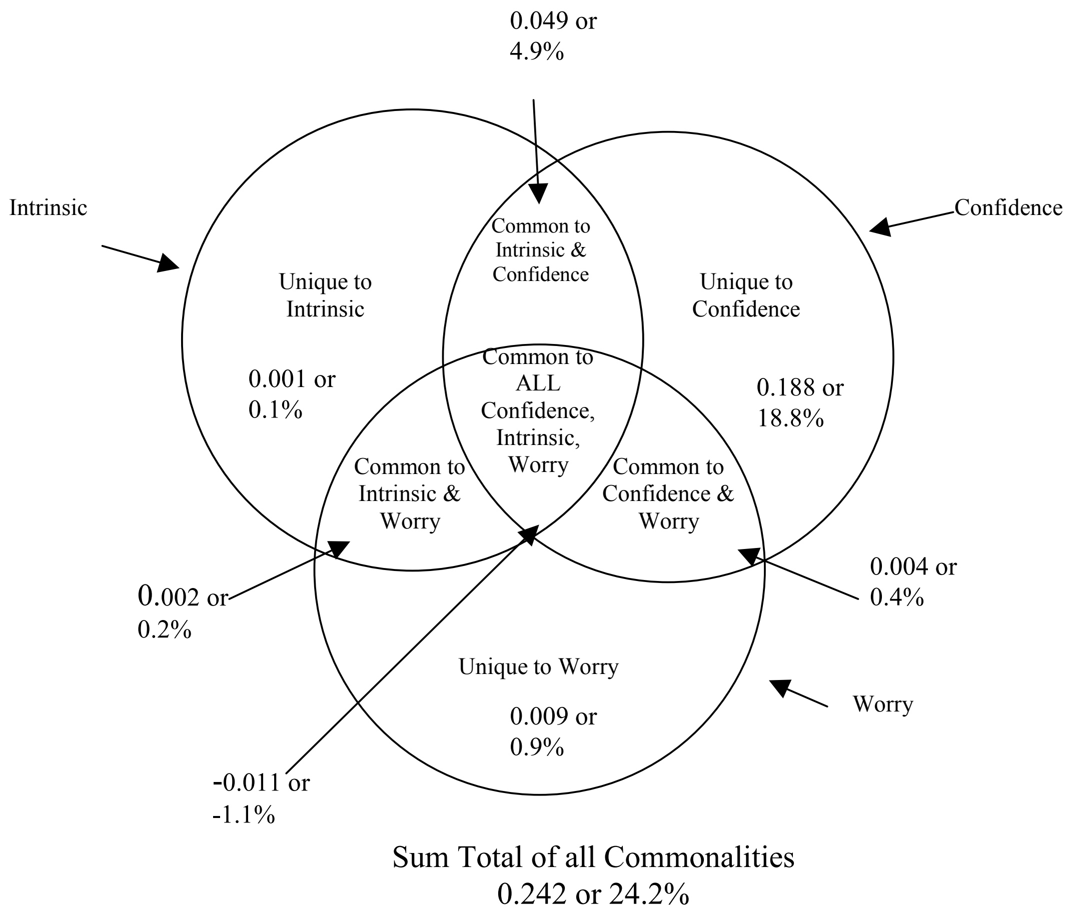

Figure 2. *Venn Diagram Showing Commonalities for Function I.*

power to partition the variance and observe which variable contributes what variance is invaluable when determining parsimony. In analyzing the data from Function I, one notices that confidence explains 18.8% of the variance alone while intrinsic value and confidence contribute 4.9 % in common. The three predictors when taken together explained 24.2% of the first function. Worry and intrinsic value explain very little of the variance from Function I, either uniquely (0.9% or 0.1%) or in common (0.1% to 0.4%) with other measured predictor variables.

Frederick (1999) stated that negative commonalities should be interpreted as zero. While Beaton (1973) believed that negative commonalities were actually confounding, increasing the predictive ability. Caution needs to be taken when interpreting the negative commonality in the common to all variables (Figure 2). As stated before in the analogy to the athlete, a negative commonality on one variable may improve the overall prediction power. However, in this case it is more appropriate to interpret the negative commonality as zero. Think of the situation this way, the variance explained by all three variables inversely predicts the variance explained when all the variables are taken separately. This scenario makes little sense and implies that the variables as a whole indicate an inverse relationship to the criterion variables where they imply a direct relationship when considered individually.

In Function I, summing the variance explained from each of the unique variables and each of the common contributions yields 0.242. The 0.242 is the variance explained in the multiple regression (R^2) and the canonical correlation Rc^2. Because CA yields the partitioned values, one would expect that the sum of the values would equal the total variance explained by either the univariate or the multivariate approach. This also illustrates that CCA subsumes the univariate case.

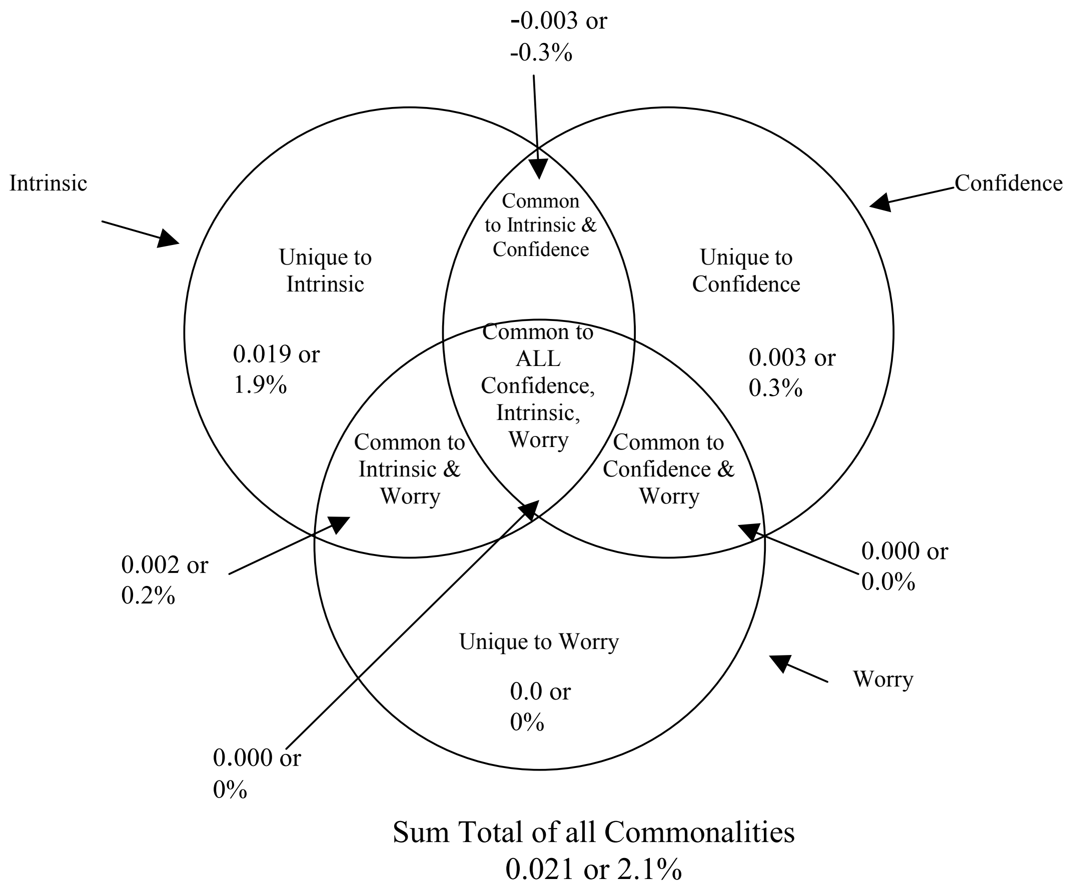

Figure 3. *Venn Diagram Showing Commonalities for Function II.*

In Function II the total variance explained is a paltry 2.1% This is hardly worthy of discussion except for the relatively large sample size to variable ratio and effect size originally indicated in the CCA. The effect size of 0.38, considered large in regards to educational research stands out in this case as well. The practical importance can not be neglected either. In review of other research on this topic, the effect size of 0.38 is large by comparison. The variance explained was partitioned into unique and common contributions and a few interesting observations are noticed.

On Function II (Figure 3) the results appear a little more interesting. Intrinsic value contributes the most variance explained 1.9% alone and confidence contributes 0.3% alone. When considering the common variance between confidence and intrinsic a -0.3% variance explained exists. This confounding seems to indicate that as the scores on confidence decreases (indicating less confidence) success on the criterion variables increase. In this case scale may influence the negative commonality. This interpretation defies logic and again implores the interpretation offered by Frederick (1999) that it should be interpreted as zero. Again, in Function II (Figure 3) worry, traditionally attributed as a major cause of poor performance in mathematics, was found to have virtually no influence.

Summary

After performing the CCA, sufficient evidence existed (i.e., an interpretable Rc^2) to continue and determine the unique and common contributions of the predictor variables. Particularly, the full model effect size of 0.38 aided the researcher in deciding to continue with further analysis. The CA yielded results on two functions. On Function I, the unique variance accounted for largely resides with the confidence variable (18.8%). This represents the overwhelming portion of the total variance 24.2% accounted for by all three of the variables - confidence, worry, and intrinsic value. This leads to an interesting supposition. First, contrary to contemporary findings this study seems to indicate that worry, contributing less than 1% of the variance, also referred to as math anxiety, is not a powerful predictor of mathematics achievement. Perhaps more time spent working on confidence and building "mathematics self-esteem" will improve mathematics achievement. Second, the results of Function II indicate that all three variables account for slightly more than 2.0% of the variance in the criterion set. This result is not very promising. However, of the variance accounted for intrinsic value accounts for 1.9 %, confidence accounts for 0.3%, and worry accounts for 0.0% of the total variance. On function II intrinsic value appears to be more helpful in predicting geometry achievement than either of the other two subscales. A list of all the SPSS syntax used in this analysis is listed in the Appendix.

The value of CA resides in the fact that the procedure yields unique and common variance explained from each of the predictor variables. The variance explained is not summative nor is it a result of interaction effects. The variance explained from the full model can be understood and the contributions of each separate variable can be interpreted in relation to the full model for the results of the unique effects. This helps to determine the most parsimonious model and relevant data sources, particularly when using a test containing subscales.

References

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

Beaton, A. E. (March, 1973). *Commonality.* (ERIC Document Reproduction Service No. ED 111 829)

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1973). *Differential aptitude tests: Administrator's handbook.* New York: Psychological Corporation

Carroll, W. M. (1998). Geometric knowledge of middle school students in a reform based mathematics curriculum. *School Science and Mathematics, 98* (4), 188-198.

DeVito, P. J. (May, 1976). The use of commonality analysis in educational research. Paper presented at the annual meeting of the New England Educational Research Association, Provincetown, MA. (ERIC Document Reproduction Service No. ED 146 210)

Fan, X. (1996). Canonical correlation analysis as a general analytic model. In B. Thompson (Ed.), *Advances in social sciences 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.

Frederick, B. N. (1999). Partitioning variance in the multivariate case: A step-by-step guide to canonical commonality analysis. In B. Thompson (Ed.), *Advances in social sciences methodology* (Vol. 5, pp. 305-318). Stamford, CT: JAI Press.

Gierl, M., & Bisanz, J. (1997). Anxieties and attitudes related to mathematics in grades 3 and 6. *Journal of Experimental Education, 63* (2), 139-158.

Henson, R. (2000). Demystifying parametric analyses: Illustrating canonical correlation analysis as the multivariate general linear model. *Multiple Linear Regression Viewpoints, 26* (1), 11-19.

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

Newton, R. G., & Spurell, D. J. (1967). Examples of the use of elements for classifying regression analysis. *Applied Statistics, 16*,165-172.

Pedhazur, E. (1982). *Multiple regression in behavior research: Explanation and prediction.* (2nd ed.). New York, NY: Holt Rinehart, and Winston.

Pedhazur, E. (1997). Canonical and discriminant analysis, and multivariate analysis of variance. In *Multiple Regression in Behavior Research.* (3rd ed., pp. 924-979). Fort Worth, TX: Harcourt Brace.

Shavelson, R. J. (1988). *Statistical reasoning for the behavioral sciences.* Boston: Allyn & Bacon.

Stevens, J. (1999). Canonical correlations. In *Applied Multivariate Statistics for the Social Sciences* (3rd ed., pp.429-449). Mahwah, NJ: Lawrence Earlbaum Associates.

Tatsuoka, M. (1971). Discriminant analysis and canonical correlation. In *Multivariate Analysis: Techniques for Educational and Psychological Research* (Chapter 6, pp. 157-194). New York: Wiley.

Thompson, B. (1984). *Canonical correlation analysis uses and interpretations*. Newbury Park, CA: Sage.

Thompson, B. (1991). Methods, plainly speaking: A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development, 24* (2), 80-93.

Thompson, B., & Miller, J. H. (February, 1985). *A multivariate method of commonality analysis*. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 263 151)

Send correspondence to: Robert M. Capraro, College of Education, Department of Teaching, Learning, & Culture, Texas A & M University, 308 Harrington Tower, College Station, TX 77843-4232. e-mail: rcapraro@coe.tamu.edu

Appendeix
SPSS Syntax for Conducting CA

Opens the file containing the data for the analyis

```
GET FILE
 "C:\WINDOWS\DESKTOP\Dissertation Data\Modified Dissertation Data File.sav"
 EXECUTE
```

Runs the descriptives that will be necessary for creating CRIT1 and CRIT2

```
 DESCRIPTIVES
   VARIABLES=spacerel gcksum
   /STATISTICS=MEAN STDDEV MIN MAX .
```

The full CCA syntax supplies the Rc^2 and the structure & function coefficients

```
Manova
 spacerel gcksum with int.val worry confid
 /print=signif(multiv eigen dimenr)
 /discrim(stan estim cor)alpha(.999))/design.
```

The syntax to create CRIT1 and CRIT2

```
COMPUTE crit1 = (.482*zspace)+(.645*zgck) .
 EXECUTE .
 COMPUTE crit2 = (-1.113*zspace)+(1.027*zgck) .
 EXECUTE .
```

All the syntax to run all possible combinations multiple regressions for the 3 predictor variables.

```
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter int.val worry confid.
 regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter int.val worry confid.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter int.val confid.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter int.val confid.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter int.val worry.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter int.val worry.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter confid worry.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter confid worry.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter int.val.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter int.val.
 regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter confid.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter confid.
regression variables=crit1 crit2 int.val worry confid/
 dependent=crit1/enter worry.
 regression variables=crit1 crit2 int.val worry confid/
 dependent=crit2/enter worry
```