Bigger is Not Better: Seeking Parsimony in Canonical Correlation Analysis via Variable Deletion Strategies

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

This paper illustrates the value of applying the law of parsimony to canonical correlation analysis (CCA) solutions. The primary purpose of parsimony is that the more parsimonious the solution, the more replicable the model will be. The ultimate goal is to estimate an equal or reasonable amount of variance with the smallest variable set possible. A real-world data set is used that is composed of 287 sixth-grade students who were administered a geometry content knowledge test with three levels and a spatial visualization test as criterion variables, and a mathematics attitude survey with six subscales as predictor variables. Three different deletion methods are delineated in the paper that will assist the researcher in deleting predictor or criterion variables to obtain a more parsimonious canonical solution.

In research contents, the law of parsimony states that the fewer variables used to explain a situation, the more probable that the explanation will be closer to reality. In a canonical correlation analysis (CCA), Thorndike (1978) stated that "as the number of variables increase, the probable effect of these sources of error variation on the canonical correlation increases" (p. 188). This is because one source of sampling error comes from the number of measured variables. Therefore, as variable sets become more parsimonious there are greater probabilities that the results of the analysis will be replicable (Cantrell, 1999).

Rim (1972) suggested that models that are more parsimonious are not only more stable and replicable but also more generalizable. According to Thompson (1989), reducing the number of variables lessen Type II error probability since degrees of freedom model are also lessened. In an analysis with three criterion variables and six predictor variables, the 18 degrees of freedom would be reduced by nine if three predictor variables were deleted from the final model. Thompson (1984a) also suggested that dropping of variables in CCA would be synonymous with "backward elimination" stepwise procedures in multiple regression. Also purported was that this connection helped to reinforce the concept that all parametric techniques are subsumed under CCA as the classical form of the general linear model (Henson, 2000; Knapp, 1976). Therefore, the goal of a variable deletion strategy is to estimate as much variance with the smallest variable set possible. This paper will show that "bigger is not better", at least in reference to the number of variables, when using canonical correlation analysis.

Since Knapp (1978) demonstrated that canonical correlation analysis was the most general form of the general linear model, CCA has gained more in popularity. Thompson (1991) showed that CCA subsumes all other parametric methods including t-tests, point bisereal, ANOVA, regression, discriminant analysis, and MANOVA. CCA has been hibernating since Hotelling first developed the logic of CCA in 1936 more than 63 years ago. Besides Knapp's demonstration, computer statistical packages have made its use more easily accessible to researchers. As Pedhazur (1997) has noted, canonical correlation matrix computation can become "prohibitive" and "complex". Modern statistical packages almost eliminate the need to create these matrixes.

Because reality involves multiple effects and multiple effects have multiple causes, canonical analysis can more accurately represents this reality by explaining multiple relationships (Clark, 1975; Thompson, 1984a). Canonical correlation analysis appropriately examines the relationship between two sets of measured variables. An example would be comparing subtests of the WISC-R and the Woodcock Johnson that measure different intellectual abilities (Eastbrook, 1984). Multiple regression analysis could do the job f there were only one dependent varaible; however, canonical analysis goes a step farther by allowing multiple dependent variables. Furthermore, CCA maximizes a set of multiplicative weights all variables in the dependent and independent variable sets (Henson, 2000). Although it is not obvious, even in multiple regression a weight is developed for the dependent variable. However, since the dependent variable is not transformed to maximize some criterion, the weight is inescapably one (1).

Table 1. Initial Solution with Canonical Communality Coefficients Deletion Strategy #I

Variable Statistic	Function 1			Function 2			Function 3			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.5	-0.845	71.40%	0.556	0.162	2.62%	0.956	0.509	25.91%	99.94%
level0	-0.179	-0.604	36.48%	1.008	0.510	26.01%	-0.617	-0.613	37.58%	100.07%
gcksum	-0.521	-0.901	81.18%	-1.197	-0.331	10.96%	-0.843	-0.279	7.78%	99.92%
Adequacy			63.02%			13.20%			23.76%	
Rd			16.13%			0.49%			0.45%	
Rc^2			25.60%			3.70%			1.9%	
*R*d			6.86%			0.68%			0.20%	
Adequacy			26.80%			18.35%			10.71%	
Useful	0.157	0.581	33.76%	0.153	-0.076	0.58%	-0.565	-0.463	21.44%	55.77%
Intrinsi	-0.096	0.426	18.15%	-0.579	-0.63	39.69%	-0.862	-0.571	32.60%	90.44%
Worry	-0.187	-0.081	0.66%	-0.829	-0.805	64.80%	0.531	0.292	8.53%	73.99%
Confid	0.932	0.972	94.48%	-0.023	-0.207	4.28%	0.787	0.083	0.69%	99.45%
Percep	0.046	0.244	5.95%	0.2	-0.061	0.37%	0.145	0.033	0.11%	6.43%
Success	0.061	0.279	7.78%	0.229	-0.061	0.37%	-0.222	-0.096	0.92%	9.08%

This present paper will illustrate three variable deletion strategies in CCA to yield the most parsimonious variable set. Parsimony will be sought for the predictor variable set, students' attitude toward mathematics, as opposed to the criterion variables, students' geometric and spatial visualization abilities. However, the same procedures could be applied to the criterion variable set.

The current data set comes from a study of 287 sixth-grade students from a south central state who were administered three tests. The *Mathematics Attitude Survey* (MATS) (Gierl & Bisanz, 1997), a Likert–type instrument, consisted of the six subscales of usefulness, intrinsic value, worry, confidence, perceptions, and attitude toward success. The six subscales served as the predictor set. The *Space Relations Portion* of the *Differential Aptitude Test* (Bennett, Seashore, & Wesman, 1973) assessed students' spatial sense focusing on visualization. The *Geometry Content Knowledge Test* (Carroll, 1998) was used to assess geometric content knowledge and to assign van Hiele (1984) geometry levels ranging from level 0 to level 2. The preceding two mathematics tests along with level 0 of the geometry content knowledge test served as the three criterion variables (spacerel, level 0, gcksum) in the study. The six subscales of the attitude survey (useful, intrinsi, worry, confid, percep, and success) served as the predictor set. The Statistical Package for the Social Sciences (SPSS) command syntax for running the CCA analysis was:

```
MANOVA
SPACEREL LEVEL0 GCKSUM WITH
USEFUL INTRINSI WORRY CONFID PERCEP SUCCESS
/PRINT=SIGNIF (MULTIV EIGEN DIMENR)
/DISCRIM (STAN ESTIM COR) ALPHA (.999)) / DESIGN.
```

The results of the analysis are compiled in Table 1, which is the suggested format for reporting canonical results.

According to Humphries-Wadsworth (1998), canonical correlation analysis is a "rich tool for examining the multiple dimensions of the synthetic variable relationships" (*p*. 6). In addition to the standardized function coefficients and structure coefficients, three other coefficients are often examined and can facilitate interpretation: canonical communality coefficients, canonical adequacy coefficients, and canonical redundancy coefficients (however, see Robert [1999] for discussion of the inadequacies of redundancy coefficients.

The researcher will now attempt to develop a clear process for completing the table. The "Func" (canonical function coefficient), the " r_s" (canonical structure coefficient) along with the Rc^2 (squared canonical correlation coefficient) for each function was obtained directly from the SPSS printout. The r_s^2 (squared canonical structure coefficient) was calculated by squaring the canonical structure coefficients for each variable and converting them into percentage format. The h^2 (communality coefficient) for each variable was obtained by summing all the r_s^2s. The adequacy coefficient, "how will a canonical variate represents the variance of the original variables in a domain" (Thompson. 1980, *p*.10), was an average of all the squared structure coefficients for the variables in one set with respect to one function. The adequacy coefficient for the criterion variable set was calculated by adding all the structure coefficients in the criterion set and dividing by the number of variables in the set and converting it into percentage format. The adequacy coefficient for the predictor set was determined by the same method. The redundancy coefficient, the redundancy of *C* (criterion variable set) given *P* (predictor variable set), was calculated by multiplying the adequacy coefficient by the Rc^2 for each function (Roberts, 1999).

After examining the full canonical analysis, the law of parsimony (Thorndike, 1978) can be invoked through a process called variable deletion. Various researchers (Cantrell, 1999; Rim, 1972; Stephens, 1996; & Thompson, 1984b) discussed approaches to achieve the most parsimonious variable set. This researcher will attempt to make the deletion process as understandable as possible. Three different strategies will be examined.

Variable Deletion

During the deletion process three coefficients will be consulted:

r_s^2 - squared canonical structure coefficient - how much variance a variable linearly shares with a canonical variate (Thompson, 1980).

$\mathbf{h^2}$ – canonical communality coefficients - sum of all $r_s^{\mathbf{2}}$; how much of the variance in a given observed variable is reproduced by the complete canonical solution (Thompson, 1991).

$R\mathrm{c}^{\mathbf{2}}$ - squared canonical coefficient– how much each function is contributing to the overall canonical solution (Thompson, 1991).

Variable Deletion Strategy #1

Deletion Strategy #1 looked at the h^2s only. The process involved the following steps:

1. Look at all the h^2s
2. Find the lowest h^2 and delete the corresponding variable
3. Rerun the CCA and recalculate the h^2s
4. Check the change to the Rc^2 for each function
5. If there is little change to Rc^2 find the next lowest h^2
6. Delete the variable with the corresponding lowest h^2 and repeat the process until the Rc^2 change is too great by researcher judgment.

Looking at Table 1, the predictor variables with the lowest h^2s were perceptions (6.34%) and success (9.08%). Both of these variables were quite a bit lower than the other four-predictor variables that ranged from 55.77 % to 99.45%. Through variable deletion strategy #1, the variable with the lowest h^2, perceptions, was dropped first. Table 2 showed the canonical analysis after perceptions was dropped. The Rc^2s were then examined for each function and there was only a very slight change. Function 1 did not change, Function 2 went from 3.7% to 3.6 %, and Function 3 remained the same. The Rc^2 change was less than 0.2% for only one function.

The remaining canonical solution still contained success with a h^2 of 9.0%. That variable was considerably lower than the other variables in Table 2, therefore, success was dropped and little change (less than 0.2%) was seen in the Rc^2s of each function as shown in Table 3. Function 1 changed from 25.6% to 25.5%, Function 2 changed from 3.6 % to 3.4%, and Function 3 changed from 1.9% to 1.8%. The limitations to this strategy involved the contributions that were not evaluated until after the variable was dropped. This could have caused keeping a large h^2 that only happened on the last canonical function and had a small Rc^2 effect size. Despite these limitations, the goal of parsimony was achieved by removing the two variables and only a very small change was noted in either the communality coefficients or the squared canonical coefficients of each function.

Table 2. *Canonical Solution After Dropping Perceptions Based on Canonical Communality Coefficients Deletion Strategy #I, Iteration #2*

Variable Statistic	Function 1			Function 2			Function 3			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.503	-0.846	71.57%	0.522	0.142	2.02%	0.974	0.513	26.32%	99.90%
level0	-0.181	-0.605	36.60%	1.028	0.528	27.88%	-0.583	-0.596	35.52%	100.00%
gcksum	-0.516	-0.9	81.00%	-1.181	-0.324	10.50%	-0.524	-0.292	8.53%	100.02%
Adequacy			63.06%			13.46%			23.45%	
Rd			16.14%			0.48%			0.45%	
Rc^2			25.60%			3.6%			1.9%	
*R*d			6.62%			0.67 %			0.21%	
Adequacy			25.85%			18.60%			11.04%	
Useful	0.167	0.581	33.76%	0.211	-0.061	0.37%	-0.53	-0.467	21.81%	55.94%
Intrinsi	-0.093	0.427	18.23%	-0.56	-0.622	38.69%	-0.891	-0.603	36.36%	93.28%
Worry	-0.177	-0.08	0.64%	-0.817	-0.825	68.06%	0.525	0.255	6.50%	75.21%
Confid	0.934	0.973	94.67%	-0.03	-0.204	4.16%	0.802	0.079	0.62%	99.46%
Percep	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%
Success	0.072	0.279	7.78%	0.286	-0.057	0.32%	-0.176	-0.098	0.96%	9.07%

Table 3. *Final Canonical Solution After Dropping Perceptions and Success Based on Communality Coefficients Deletion Strategy #I, Iteration #3*

Variable Statistic	Function 1			Function 2			Function 3			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.504	-0.846	71.57%	0.583	0.171	2.92%	0.938	0.505	25.50%	100.00%
level0	-0.190	-0.610	37.21%	0.984	0.482	23.23%	-0.651	-0.628	39.44%	99.88%
gcksum	-0.509	-0.898	80.64%	-1.218	-0.349	12.18%	-0.441	-0.266	7.08%	99.90%
Adequacy			63.14%			12.78%			24.01%	
Rd			16.10%			0.43%			0.43%	
Rc^2			25.50%			3.40%			1.80%	
*R*d			6.29%			0.67%			0.20%	
Adequacy			24.67%			19.78%			10.97%	
Useful	0.175	0.582	33.87%	0.229	-0.075	0.56%	-0.584	-0.475	22.56%	57.00%
Intrinsi	-0.093	0.43	18.49%	-0.629	-0.664	44.09%	-0.845	-0.557	31.02%	93.60%
Worry	-0.153	-0.078	0.61%	-0.732	-0.840	70.56%	0.549	0.339	11.49%	82.66%
Confid	0.950	0.975	95.06%	0.082	-0.187	3.50%	0.764	0.087	0.76%	99.32%
Percep	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%

Variable Deletion #2

Deletion Strategy #2 looks at the contribution of each function the total canonical solution. The steps in the process are as follows:

1. Run a full CCA and look at the Rc^2 for each function.
2. Omit the function with the smallest Rc^2
3. Compute the subset of h^2s
4. Now find variable that has the lowest h^2; drop it from the original solution
5. Repeat the process until the remaining variables are reasonably close in their subset h^2 values. This will be a matter of researcher judgment.

The researcher employed strategy #2 in order to consider the value of each function to the whole canonical solution. Looking at Table 1, the lowest squared canonical coefficient (Rc^2) was found in Function 3 (1.9%), thus the entire function was dropped (Table 4). Note that the h^2 still showed that the

Table 4. *InitialCanonical Solution After Dropping Function 3 with Subset Canonical Communality Coefficients Deletion Strategy #2, Iteration #1*

Variable Statistic	Function 1			Function 2			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.5	-0.845	71.40%	0.556	0.162	2.62%	74.03%
level0	-0.179	-0.604	36.48%	1.008	0.510	26.01%	62.49%
gcksum	-0.521	-0.901	81.18%	-1.197	-0.331	10.96%	92.14%
Adequacy			63.02%			13.20%	
Rd			16.13%			0.49%	
Rc^2			25.60%			3.70%	
*R*d			6.86%			0.68%	
Adequacy			26.80%			18.35%	
Useful	0.157	0.581	33.76%	0.153	-0.076	0.58%	34.33%
Intrinsi	-0.096	0.426	18.15%	-0.579	-0.63	39.69%	57.84%
Worry	-0.187	-0.081	0.66%	-0.829	-0.805	64.80%	65.46%
Confid	0.932	0.972	94.48%	-0.023	-0.207	4.28%	98.76%
Percep	0.046	0.244	5.95%	0.2	-0.061	0.37%	6.33%
Success	0.061	0.279	7.78%	0.229	-0.061	0.37%	8.16%

Table 5. *Canonical Solution After Dropping Perceptions Based on Subset Canonical Communality Coefficients Deletion Strategy #2, Iteration #2*

Variable Statistic	Function 1			Function 2			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.503	-0.846	71.57%	0.522	0.142	2.02%	73.59%
level0	-0.181	-0.605	36.60%	1.028	0.528	27.88%	64.48%
gcksum	-0.516	-0.9	81.00%	-1.181	-0.324	10.50%	91.50%
Adequacy			63.06%			13.46%	
Rd			16.14%			0.48%	
Rc^2			25.60%			3.60%	
*R*d			6.62%			0.67%	
Adequacy			25.85%			18.60%	
Useful	0.167	0.581	33.76%	0.211	-0.061	0.37%	34.13%
Intrinsi	-0.093	0.427	18.23%	-0.56	-0.622	38.69%	56.92%
Worry	-0.177	-0.08	0.64%	-0.817	-0.825	68.06%	68.70%
Confid	0.934	0.973	94.67%	-0.03	-0.204	4.16%	98.83%
Percep	0	0	0.00%	0	0	0.00%	0.00%
Success	0.072	0.279	7.78%	-0.286	-0.057	0.32%	8.11%

Function 3 (1.9%), thus the entire function was dropped (Table 4). Note that the h^2 still showed that the variables of perception and success had the lowest h^2s, 6.33% and 8.16% respectively. Perceptions was first variable deleted and the results of the canonical solution was displayed in Table 5. Table 6 indicated an even more parsimonious solution after dropping success. Since a subset with a close grouped $\underline{h^2}$ subset was sought, this researcher also dropped useful (34.43%). Table 7 showed the smallest set of variables with a relatively close range of communality coefficients. The h^2s were intrinsic (69.1%), worry (67.28%), and confidence (99.79%). Based on the literature and researcher judgment, the iteration process was ended. Of the three remaining variables, worry had a squared structure coefficient of .53% on Function1 but a 66.75% on Function 2. Reverse effects were seen for confidence that had a r_s^2 of 97.42% on Function1 but 2.37% on Function 2.One limitation of strategy #2 was that it did not consider functions with small Rc^2 values. In addition, the variations as to where h^2 values came from as shown in worry and confidence were not considered.

Table 6. *Canonical Solution After Dropping Perceptions and Success Based on Subset Canonical Communality Coefficients Deletion Strategy #2, Iteration #3*

Variable Statistic	Function 1			Function 2			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.504	-0.846	71.57%	0.583	0.171	2.92%	74.50%
level0	-0.190	-0.610	37.21%	0.984	0.482	23.23%	60.44%
gcksum	-0.509	-0.898	80.64%	-1.218	-0.349	12.18%	92.82%
Adequacy			63.14%			12.78%	
Rd			16.10%			0.43%	
Rc^2			25.50%			3.40%	
*R*d			6.29%			0.67%	
Adequacy			24.67%			19.78%	
Useful	0.175	0.582	33.87%	0.229	-0.075	0.56%	34.43%
Intrinsi	-0.093	0.430	18.49%	-0.629	-0.664	44.09%	62.58%
Worry	-0.153	-0.078	0.61%	-0.732	-0.840	70.56%	71.17%
Confid	0.950	0.975	95.06%	0.082	-0.187	3.50%	98.56%
Percep	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0.00%

Table 7. *Final Canonical Solution After Dropping Perceptions, Success and Useful Based on Canonical Communality Coefficients Deletion Strategy #2, Iteration #4*

Variable Statistic	Function 1			Function 2			
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.491	-0.837	70.06%	0.692	0.225	5.06%	75.12%
level0	-0.216	-0.629	39.56%	0.892	0.393	15.44%	55.01%
gcksum	-0.503	-0.9	81.00%	-1.268	-0.389	15.13%	96.13%
Adequacy			63.54%			11.88%	
Rd			16.20%			0.40%	
Rc^2			25.50%			3.40%	
*R*d			4.96%			0.68%	
Adequacy			19.46%			19.90%	
Useful	0	0	0.00%	0	0	0.00%	0.00%
Intrinsi	-0.065	0.434	18.84%	-0.682	-0.709	50.27%	69.10%
Worry	-0.139	-0.073	0.53%	-0.68	-0.817	66.75%	67.28%
Confid	1.031	0.987	97.42%	0.249	-0.154	2.37%	99.79%
Percep	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0.00%

Variable Deletion #3

Deletion Strategy # 3 considered weighted h^2. This strategy looked at the variables' contribution to the complete canonical solution. The steps were as follows:

1. Multiply Rc^2 by each r_s^2 and add the products together for each function to obtain the weighted h^2 for each variable.
2. Drop the lowest weighted h^2, repeat the previous step.
3. Look at the change in Rc^2; if there is little change, drop the variable with the next lowest h^2.
4. Take out as many variables as possible without compromising the Rc^2.

In order to consider the limitations of variable deletion #2, the weighted communality coefficients helped the researcher obtain a more realistic view of how much each predictor variable contributes to the total canonical analysis. Using the above algorithm in step 1, the weighted communality coefficients

Table 8. *Initial Canonical Solution with Weighted Canonical Communality Coefficients Deletion Strategy #3, Iteration #1*

Variable Statistic	Function 1			Function 2			Function 3			Weighted
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.500	-0.845	71.40%	0.556	0.162	2.62%	0.956	0.509	25.91%	18.87%
level0	-0.179	-0.604	36.48%	1.008	0.510	26.01%	-0.617	-0.613	37.58%	11.02%
gcksum	-0.521	-0.901	81.18%	-1.197	-0.331	10.96%	-0.843	-0.279	7.78%	21.34%
Adequacy			63.02%			13.20%			23.76%	
Rd			16.13%			0.49%			0.45 %	
Rc^2			25.60%			3.70%			1.9%	
*R*d			6.86%			0.68%			0.20%	
Adequacy			26.80%			18.35%			10.71%	
Useful	0.157	0.581	33.76%	0.153	-0.076	0.58%	-0.565	-0.463	21.44%	9.07%
Intrinsi	-0.096	0.426	18.15%	-0.579	-0.63	39.69%	-0.862	-0.571	32.60%	6.73%
Worry	-0.187	-0.081	0.66%	-0.829	-0.805	64.80%	0.531	0.292	8.53%	2.73%
Confid	0.932	0.972	94.48%	-0.023	-0.207	4.28%	0.787	0.083	0.69%	24.36%
Percep	0.046	0.244	5.95%	0.2	-0.061	0.37%	0.145	0.033	0.11%	1.54%
Success	0.061	0.279	7.78%	0.229	-0.061	0.37%	-0.222	-0.096	0.92%	2.02%

Table 9. *Canonical Solution with Canonical Weighted Communality Coefficients After Dropping Perceptions Deletion Strategy #3, Iteration 2*

Variable Statistic	Function 1			Function 2			Function 3			Weighted
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.503	-0.846	71.57%	0.522	0.142	2.02%	0.974	0.513	26.32%	18.89%
level0	-0.181	-0.605	36.60%	1.028	0.528	27.88%	-0.583	-0.596	35.52%	11.05%
gcksum	-0.516	-0.900	81.00%	-1.181	-0.324	10.50%	-0.524	-0.292	8.53%	21.28%
Adequacy			63.06%			13.46%			23.45%	
Rd			16.14%			0.48%			0.45%	
Rc^2			25.60%			3.60%			1.90%	
*R*d			6.62%			0.67%			0.21%	
Adequacy			25.85%			18.60%			11.04%	
Useful	0.167	0.581	33.76%	0.211	-0.061	0.37%	-0.530	-0.467	21.81%	9.07%
Intrinsi	-0.093	0.427	18.23%	-0.560	-0.622	38.69%	-0.891	-0.603	36.36%	6.75%
Worry	-0.177	-0.080	0.64%	-0.817	-0.825	68.06%	0.525	0.255	6.50%	2.74%
Confid	0.934	0.973	94.67%	-0.030	-0.204	4.16%	0.802	0.079	0.62%	24.40%
Percep	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%
Success	0.072	0.279	7.78%	0.286	-0.057	0.32%	-0.176	-0.098	0.96%	2.02%

were obtained and examined. Table 8 illustrated the entire canonical solution showing weighted communality coefficients. Since the variable perceptions had the lowest weighted h^2 (1.54%), it was first dropped resulting in Table 9. The next lowest, success (2.02%), was then deleted resulting in Table 10. The next smallest weighted h^2 came from worry (2.91%), which was then deleted. The results are displayed in Table 11. After these three deletions from the canonical solution, the Rc^2 changes were small, 0.7% in Function 1, 1.3% in Function 2, and 1.4% in Function 3.

Since none of the variables remaining had their highest squared structure coefficient (r_s^2) in Function 3, which also had the lowest Rc^2 (0.5%), Function 3 was now dropped and the most parsimonious solution set resulted in two functions with three predictors displayed in Table 12. The researcher considered this the best combination of the deletion strategies since both the functions and the weighted h^2s were considered. The results indicated that when students consider mathematics useful and most importantly are confident in mathematics, they perform better on tests that measure their geometric content knowledge

Table 10. *Initial Solution with Canonical Weighted Communality Coefficients After Dropping Perceptions and Success Deletion Strategy #3, Iteration*

Variable Statistic	Function 1			Function 2			Function 3			Weighted h^2
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	
spacerel.	-0.504	-0.846	71.57%	0.583	0.171	2.92%	0.938	0.505	25.50%	18.91%
level0	-0.190	-0.610	37.21%	0.984	0.482	23.23%	-0.651	-0.628	39.44%	11.11%
gcksum	-0.509	-0.898	80.64%	-1.218	-0.349	12.18%	-0.441	-0.266	7.08%	21.22%
Adequacy			63.14%			12.78%			24.01%	
Rd			16.16%			0.46%			0.46%	
Rc^2			25.60%			3.60%			1.90%	
*R*d			6.32%			0.71%			0.21%	
Adequacy			24.67%			19.78%			10.97%	
Useful	0.175	0.582	33.87%	0.229	-0.075	0.56%	-0.584	-0.475	22.56%	9.12%
Intrinsi	-0.093	0.43	18.49%	-0.629	-0.664	44.09%	-0.845	-0.557	31.02%	6.91%
Worry	-0.153	-0.078	0.61%	-0.732	-0.840	70.56%	0.549	0.339	11.49%	2.91%
Confid	0.950	0.975	95.06%	0.082	-0.187	3.50%	0.764	0.087	0.76%	24.48%
Percep	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%

Table 11. *Final Canonical Solution After Dropping Perceptions, Success, and Worry with Weighted Canonical Communality Coefficients Deletion Strategy 3, Iteration 4*

Variable Statistic	Function 1			Function 2			Function 3			Weighted h^2
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	Func.	r_s	r_s^2	
spacerel.	-0.508	-0.841	70.73%	-1.038	0.439	19.27%	0.371	0.316	9.99%	18.12%
level0	-0.244	-0.643	41.34%	0.338	-0.041	0.17%	-1.12	-0.765	58.52%	10.59%
gcksum	-0.468	-0.889	79.03%	-1.228	-0.455	20.70%	0.459	0.056	0.31%	20.18%
Adequacy			63.70%			13.38%			22.94%	
Rd			15.86%			0.32%			0.11%	
Rc^2			24.90%			2.40%			0.50%	
*R*d			6.27%			0.35%			0.05%	
Adequacy			25.19%			14.50%			10.27%	
Useful	0.159	0.586	34.34%	-0.129	-0.294	8.64%	-1.15	-0.755	57.00%	9.04%
Intrinsi	-0.124	0.44	19.36%	-1.131	-0.883	77.97%	0.345	0.16	2.56%	6.70%
Worry	0	0	0.00%	0	0.000	0.00%	0	0	0.00%	0.00%
Confid	0.973	0.987	97.42%	0.581	-0.064	0.41%	0.529	0.144	2.07%	24.28%
Percep	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0	0	0.00%	0.00%

and spatial visualization. Also, students who receive extrinsic rewards perform better than those students who rely on intrinsic motivation.

The goal of all these deletion strategies was a more parsimonious solution. Therefore, choosing the smaller variable set when the same amount of variance can be accounted for was achieved. Just remember “bigger is not better!” in canonical correlation analysis.

Table 12. *Final Canonical Solution with Combination of Variable Deletion Strategies With Weighted Canonical Communality Coefficients*

Variable Statistic	Function 1			Function 2			Weighted
	Func.	r_s	r_s^2	Func.	r_s	r_s^2	h^2
spacerel.	-0.508	-0.841	70.73%	-1.038	0.439	19.27%	18.07%
level0	-0.244	-0.643	41.34%	0.338	-0.041	0.17%	10.30%
gcksum	-0.468	-0.889	79.03%	-1.228	-0.455	20.70%	20.18%
Adequacy			63.70%			13.38%	
Rd			15.86%			0.32%	
Rc^2			24.90%			2.40%	
*R*d			6.27%			0.35%	
Adequacy			25.19%			14.50%	
Useful	0.159	0.586	34.34%	-0.129	-0.294	8.64%	8.76%
Intrinsi	-0.124	0.44	19.36%	-1.131	-0.883	77.97%	6.69%
Worry	0	0	0.00%	0	0.000	0.00%	0.00%
Confid	0.973	0.987	97.42%	0.581	-0.064	0.41%	24.27%
Percep	0	0	0.00%	0	0	0.00%	0.00%
Success	0	0	0.00%	0	0	0.00%	0.00%

References

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

Cantrell, C. (1999). Variable deletion. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 321-333). Stamford, CT: J.A.I. Press.

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

Clark, D. (1975). *Understanding canonical correlation analysis*. Norwich : Geo Abstracts Ltd.

Estabrook, G. E. (1984). A canonical correlation analysis of the wechsler intelligence scale for children-revised and the woodcock-johnson tests of cognitive ability in a sample referred for suspected learning disabilities. *Journal of Educational Psychology, 76* (6), 1170-1177.

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

Humphries-Wadsworth, T. (1998, April). *Features of published analyses of canonical results.* Paper presented at annual meeting of the American Educational Research Association, San Diego, CA. (ERIC Document Reproduction Service No. ED 418 125)

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

Pedhazur, E. J. (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.

Rim, E. (1975, April). *A stepwise canonical procedure and the shrinkage of canonical correlations.* Paper presented at annual meeting of the American Educational Research Association, Washington, DC. (ERIC Document Reproduction Service No. ED 109 237)

Roberts, J.K. (1999). Canonical redundancy (Rd) coefficients they should (almost never) be computed and interpreted. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 331-333). Stamford, CT: J.A.I.Press.

Stevens, J. (1999). Canonical correlations In *Applied multivariate statistics for the social sciences* (3rd ed., pp. 429-449). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Thompson, B. (1982, February). *Stepwise canonical correlation analysis: A new research technique*. Paper presented at annual meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 222 546)

Thompson, B. (1984a). *Canonical correlation analysis: Uses and interpretations*. Beverly Hills, CA: Sage.

Thompson, B. (1984b, April). *Canonical correlation analysis: An annotated bibliography*. Paper presented at annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 242 792)

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.

Thorndike, R. (1978). *Correlation procedures for research.* New York: Gardner.

van Hiele-Geldorf, D. (1984). The didactics of geometry in the lowest class of secondary school. In D. Fuys, D. Geddes, & R. Tischer (Eds.), *English translation of selected writings of Dina van Hiele-Geldorf and Pierre M. van Hiele* (pp. 1-214). Brooklyn, NY: Brooklyn College, School of Education. (ERIC Document Reproduction Service No. 289 697)

van Hiele, P.M. (1987, June). *A method to facilitate the finding of levels of thinking in geometry by using the levels in arithmetic*. Paper presented at the working conference for Learning and Teaching Geometry: Issues for Research and Practice, Syracuse, NY.

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

---