Journal of Methods and Measurement in the Social Sciences,
Vol. 6, No. 2, 30-41, 2015

The Case for Using the General Linear Model as a Unifying Conceptual Framework for Teaching Statistics and Psychometric Theory

Bruce Thompson
Texas A&M University and Baylor College of Medicine

The present paper argues for teaching statistics and psychometric theory using the GLM as a unifying conceptual framework. This helps students understand what analyses have in common, and also provides a firm grounding for understanding that more general cases of the GLM (canonical correlation analysis and SEM) can be interpreted with the same rubric used throughout the GLM. And this approach also helps students better understand analyses that are *not* part of the GLM, such as predictive discriminant analysis (PDA). The approach helps students understand that *all* GLM analyses (a) are correlational, and thus are *all* susceptible to sampling error, (b) can yield r^2-type effect sizes, and (c) use weights applied to measured variables to estimate the latent variables really of primary interest.

Keywords: General Linear Model, significance testing, effect sizes

In a very influential APA presidential address in the late 1950s, Lee Cronbach advocated greater use of aptitude-treatment interaction designs, and effectively decried the then common misconception that statistics could be conceptualized as fitting within two classes: experimental statistics and correlational statistics. But not much happened with respect to how researchers conceptualized and taught statistics and psychometrics.

Then, in 1968, Cohen published a seminal article that was almost as important as his 1994 article, "The Earth is round ($p<.05$)." In the 1968 article, Cohen said, although he thought most statisticians would find his argument obvious, most psychologists at the time on the other hand had no idea that regression subsumed ANOVA and other univariate analyses as special cases. Thus, regression is the univariate general linear model (GLM). He argued that the GLM was important conceptually, but also that very important advantages could be realized by using regression to conduct many univariate analyses.

Subsequently, Knapp (1978) showed that canonical correlation analysis (CCA; see Thompson, 1984, 2000) was the multivariate GLM, subsuming in addition to other multivariate methods (e.g., Hotelling T^2, descriptive discriminant analysis [but not predictive discriminant analysis], and MANOVA and MANCOVA) univariate regression and the other univariate parametric methods. Finally, Bagozzi, Fornell, and Larcker (1981; also see Fan, 1997) showed that structural equation modeling (SEM) was the most general case of the GLM.

A virtual regression-discontinuity study of the influence of Cohen's article shows that the field changed dramatically following Cohen's (1968) publication. Studies by Edgington (1964, 1974) covering several decades showed that prior to the 1968 article around 2/3rds to 3/4ths of published articles used ANOVAs. Similar studies after the 1968 article showed a large drop in the use of ANOVAs (Elmore & Woehlke, 1988; Goodwin & Goodwin, 1985; Kieffer, Reese & Thompson, 2001; Willson, 1980).

The present paper argues for teaching statistics and psychometric theory using the GLM as a unifying conceptual framework. This helps students understand what analyses have in common, and also provides a firm grounding for understanding that more general cases of the GLM (canonical correlation analysis and SEM) can be interpreted with the same rubric used throughout the GLM. And this approach also helps students better understand analyses that are *not* part of the GLM, such as predictive discriminant analysis (PDA). The approach helps students understand that *all* GLM analyses (a) are correlational, and thus are *all* susceptible to sampling error, (b) can yield r^2-type effect sizes, and (c) use weights applied to measured variables to estimate the latent variables really of primary interest.

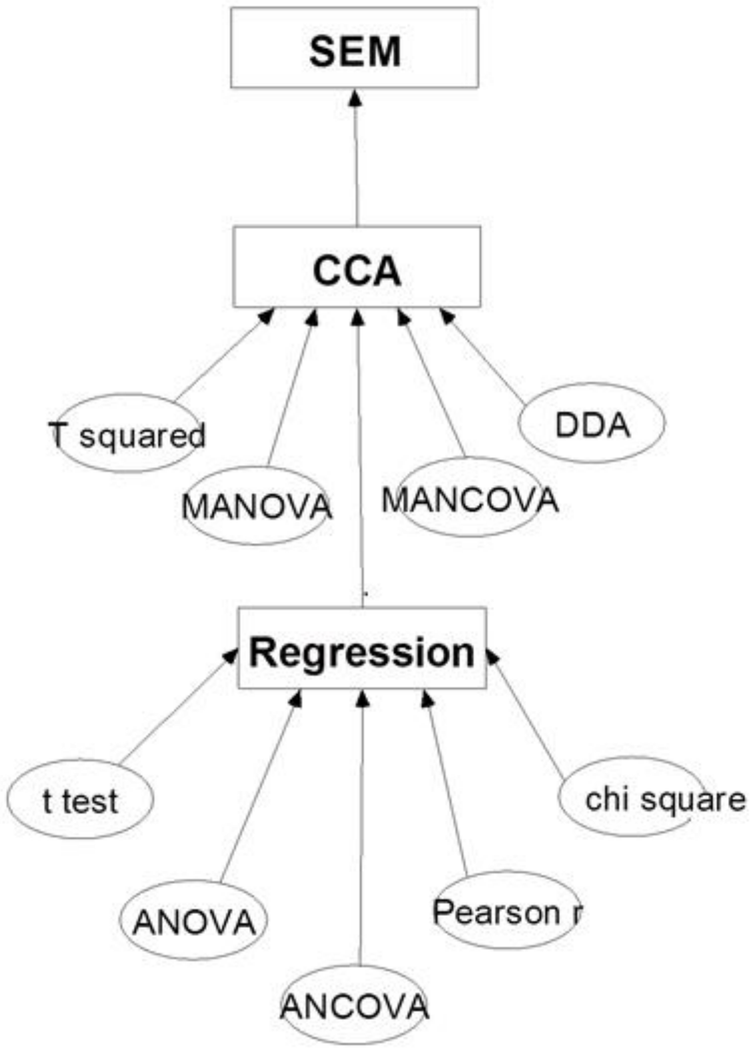
More on Defining the GLM

The General Linear Model is the concept that "all analytic methods are correlational ... and yield variance-accounted-for effect sizes analogous to (e.g., R^2, η^2, ω^2)" (Thompson, 2000, p. 263). As Graham (2008) explained,

> The vast majority of parametric statistical procedures in common use are part of [a single analytic family called] the General Linear Model (GLM), including the *t* test, analysis of variance (ANOVA), multiple regression, descriptive discriminant analysis (DDA), multivariate analysis of variance (MANOVA), canonical correlation analysis (CCA), and structural equation modeling (SEM). Moreover, these procedures are *hierarchical* [italics added], in that some procedures are special cases of others. (p. 485)

Figure 1 presents a conceptual map of the commonly used statistical analyses falling within the General Linear Model. As noted previously, predictive discriminant analysis (PDA), unlike descriptive discriminant analysis (DDA), is *not* part of the GLM (Huberty, 1994). It can also be shown that the mathematics of factor analysis are used to compute the multiplicative weights applied to the measured variables, either explicitly or implicitly, in *all* analyses throughout the GLM.

Figure 1. Conceptual Map of the General Linear Model



Note. Descriptive discriminant analysis (DDA) is part of the general linear model, but predictive discriminant analysis (PDA) is *not* part of the GLM. "SEM" = structural equation modeling; "CCA" = canonical correlation analysis; "*T*-squared" = Hotelling's T^2, a multivariate extension of the *t*-test to multiple dependent variables.

A very powerful way to prove to students that CCA subsumes other multivariate and univariate parametric methods is to use "proof by SPSS." That is, some students assume that SPSS was written by God, and that therefore anything on an SPSS output must be infallibly true. Here I use the Appendix A heuristic data and the Appendix B SPSS syntax to perform a couple of these proofs via SPSS.

CCA Subsumes Regression as a Special Case

Figure 2 presents a cut-and-paste copy out of an SPSS output for a regression analysis predicting IQ scores with predictors X1, X2, and X3. Figure 3 presents a cut-and-paste copy out of an SPSS output for a CCA with the same measured variables. Note that the regression R^2 = .04106 equals the CCA R_C = .041.

Figure 2. Cut-and-Paste Copy of SPSS (version 6) Output Regression with IQ as Outcome Variable and Variables X1, X2 and X3 as Predictors

```
* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Listwise Deletion of Missing Data

Equation Number 1    Dependent Variable..   IQ

Block Number  1.  Method:  Enter      X1       X2       X3

Variable(s) Entered on Step Number  1..   X3
                                    2..    X2
                                    3..    X1

Multiple R         .20263   Analysis of Variance
R Square          .04106                     DF   Sum of Squares   Mean Square
Adjusted R Square   -.16443   Regression          3         17.42991       5.80997
Standard Error      5.39226   Residual           14        407.07009      29.07644

                        F =        .19982       Signif F =  .8948

------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T

X1             .029676     .047711    .178198      .622  .5439
X2             .016907     .051648    .089691      .327  .7482
X3             .016575     .046738    .100762      .355  .7281
(Constant)    99.489958    4.497892              22.119  .0000
```

Figure 3. Cut-and-Paste Copy of SPSS (version 6) Output CCA with IQ as Predictor Variable and Variables X1, X2 and X3 as Outcomes

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Eigenvalues and Canonical Correlations

Root No.    Eigenvalue   Pct.     Cum. Pct.  Canon Cor.     Sq. Cor
    1          .043     100.000    100.000     .203           .041
 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Standardized canonical coefficients for DEPENDENT variables
           Function No.

 Variable           1
 X1             -.879
 X2             -.443
 X3             -.497
```

However, at first glance the regression beta weights do not appear to match the CCA standardized function coefficients for the parallel analysis. First, the CCA function coefficients each have a different sign than the three regression beta weights! But the scaling direction of equations is purely arbitrary, and any researcher can at will reverse all the signs in an equation within the GLM (see Thompson, 2004, pp. 96-97). This is the equivalent of the arbitrary choice of whether to score a test by counting number of right answers versus number of wrong answers.

Second, the scaling of the regression beta weights and the CCA function coefficients is different. Table 1 illustrates how the two sets of weights can be converted into each other's metrics.

Table 1

Converting Beta Weights into CCA Function Coefficients, and Vice Versa

Variable	Beta/R_C	Function Coefs.		R
X1	0.178198 / 0.203 =	0.879	*	0.20263
X2	0.089691 / 0.203 =	0.443	*	0.20263
X3	0.100762 / 0.203 =	0.497	*	0.20263

CCA Subsumes Multi-Way ANOVA as a Special Case

Figure 4 presents a cut-and-paste copy out of an SPSS output for an ANOVA summary table for an analysis into IQ as the dependent variable in a two-way factorial ANOVA. Conducting the parallel ANOVA using CCA is a bit

Figure 4. Cut-and-Paste Copy of SPSS (version 6) Output ANOVA with IQ as Outcome Variable and Variables X1, X2 and X3 as Predictor Variables

```
* * *  A N A L Y S I S   O F   V A R I A N C E  * * *

                 IQ
            by   EXP_GRP
                 GENDER

                 UNIQUE sums of squares
                 All effects entered simultaneously

                         Sum of                  Mean              Sig
Source of Variation               Squares     DF        Square        F    of F

Main Effects                      412.500      3       137.500   137.500  .000
   EXP_GRP                        300.000      2       150.000   150.000  .000
   GENDER                         112.500      1       112.500   112.500  .000
2-Way Interactions                   .000      2          .000      .000  1.00
   EXP_GRP  GENDER                   .000      2          .000      .000  1.00
Explained                         412.500      5        82.500    82.500  .000
Residual                           12.000     12         1.000
Total                             424.500     17        24.971
```

tedious, but otherwise is not problematic. First, create orthogonal contrasts using conventional methods explained in various textbooks (e.g., Thompson, 2006). Next, run a CCA model using all the contrast variables. Then run CCA models dropping in turn the contrast variables for the three omnibus effects. Figure 5 presents cut-and-paste copies out of an SPSS output for these analyses.

Figure 5. Cut-and-Paste Copy of SPSS (version 6) Output CCAs with IQ as IQ as Outcome Variable and Variables X1, X2 and X3 as Predictor Variables.

```
CCA Model #1 with All Five Orthogonal Contrasts Entered
   EFFECT .. WITHIN CELLS Regression
   Multivariate Tests of Significance (S = 1, M = 1 1/2, N = 5 )

   Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F
   Pillais          .97173   82.50000       5.00      12.00       .000
   Hotellings     34.37500   82.50000       5.00      12.00       .000
   Wilks            .02827   82.50000       5.00      12.00       .000
   Roys             .97173
   Note.. F statistics are exact.

CCA Model #2 Omitting Orthogonal Contrasts for the Three-Level "A" Way
   EFFECT .. WITHIN CELLS Regression
   Multivariate Tests of Significance (S = 1, M = 1/2, N = 6 )

   Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F
   Pillais          .26502    1.68269       3.00      14.00       .216
   Hotellings       .36058    1.68269       3.00      14.00       .216
   Wilks            .73498    1.68269       3.00      14.00       .216
   Roys             .26502
   Note.. F statistics are exact.

CCA Model #3 Omitting the Orthogonal Contrast for the Two-Level "B" Way
   EFFECT .. WITHIN CELLS Regression
   Multivariate Tests of Significance (S = 1, M = 1 , N = 5 1/2)

   Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F
   Pillais          .70671    7.83133       4.00      13.00       .002
   Hotellings      2.40964    7.83133       4.00      13.00       .002
   Wilks            .29329    7.83133       4.00      13.00       .002
   Roys             .70671
   Note.. F statistics are exact.

CCA Model #4 Omitting the Orthogonal Contrasts for the Interaction Effects
   EFFECT .. WITHIN CELLS Regression
   Multivariate Tests of Significance (S = 1, M = 1/2, N = 6 )

   Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F
   Pillais          .97173  160.41667       3.00      14.00       .000
   Hotellings     34.37500  160.41667       3.00      14.00       .000
   Wilks            .02827  160.41667       3.00      14.00       .000
   Roys             .97173
   Note.. F statistics are exact.
```

Tables 2 and 3 present the calculations to convert CCA lambda values back into conventional ANOVA $F_{\text{CALCULATED}}$ values. Obviously, one would not routinely perform ANOVA using CCA, but nevertheless the point that CCA is the multivariate GLM has hopefully been made!

Table 2
Step #1 in Converting CCA Results into Conventional ANOVA $F_{CALCULATED}$ Values: Running Full and Restricted Models for the Various ANOVA Omnibus Effects

	Model	Predictors				CCA lambda
1	A1	A2	B1	A1B1	A2B1	0.02827
2			B1	A1B1	A2B1	0.73498
3	A1	A2		A1B1	A2B1	0.29329
4	A1	A2	B2			0.02827

Table 3
Step #2: Converting CCA lambda into Ratios into Classical ANOVA $F_{CALCULATED}$ Values

Effect	Ratio	Full Mode lambda	Lambda w/o effect	Ratio
A Way	1 / 2	0.02827	0.73498	0.03846
B Way	1 / 3	0.02827	0.29329	0.09639
AxB Interaction	1 / 4	0.02827	0.02827	1.00000

Reasons Why I Advocate Teaching Statistics and Psychometrics from a GLM Perspective

Here are some of the reasons why I advocate teaching both statistics and psychometrics from a GLM perspective:

1. Teaching statistics and psychometrics from a GLM perspective helps students understand that sampling error effects both $p_{\text{CALCULATED}}$ values and effect sizes no matter what analysis is being done.
2. Teaching statistics and psychometrics from a GLM perspective helps students understand that weights are applied to measured variables to estimate latent variables in every analysis. For example, in a balanced ANOVA the eta values are also the beta weights for estimating Y-hat values in ANOVA (see Thompson, 2006).
3. Teaching statistics and psychometrics from a GLM perspective helps students understand that it is the design, and *not* the analysis, that

provides the ability to make causal claims (see Thompson, 2006, chapter 12).

4. Teaching statistics and psychometrics from a GLM perspective helps students understand that statistics and psychometric models do the same things, albeit it for different purposes: they partition variances (or sum of squares) and estimate ratios of those partitions in forms such as eta squared, R_2, R_{C2}, and reliability coefficients (see Dawson, 1999).
5. Teaching statistics and psychometrics from a GLM perspective helps students understand that all analyses are correlation, which among other things implies that all analyses can be conducted without a researcher's data, as long as one has the covariance matrix and means and *SD*s. These summary statistics are perfectly suitable as inputs into SPSS analyses (see Zientek & Thompson, 2009).

Author Notes: Correspondence concerning this article should be addressed to Bruce Thompson at bruce-thompson@tamu.edu.

References

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.

Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin, 70*, 426-433.

Dawson, T.E. (1999). Relating variance partitioning in measurement analyses to the exact same process in substantive analyses. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 101-110). Stamford, CT: JAI Press.

Edgington, E.S. (1964). A tabulation of inferential statistics used in psychology journals. *American Psychologist, 19*, 202-203.

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

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.

Graham, J. M. (2008). The General Linear Model as structural equation modeling. *Journal of Educational and Behavioral Statistics, 33*, 485-506.

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

Kieffer, K. M., Reese, R. J., & Thompson, B. (2001). Statistical techniques employed in *AERJ* and *JCP* articles from 1988 to 1997: A methodological review. *Journal of Experimental Education, 69*, 280-309.

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

Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretations*. Thousand Oaks, CA: Sage.

Thompson, B. (2000). Canonical correlation analysis. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 285-316). Washington, DC: American Psychological Association.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach*. New York: Guilford.

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

Zientek, L.R., & Thompson, B. (2009). Matrix summaries improve research reports: Secondary analyses using published literature. *Educational Researcher, 38*, 343-352.

Appendix A
Hypothetical Heuristic Data

1	1	1	94	790	12	41	14
2	1	1	95	795	51	6	88
3	1	1	96	800	79	99	32
4	1	2	99	780	9	16	79
5	1	2	100	785	60	33	14
6	1	2	101	790	97	30	3
7	2	1	99	785	28	64	90
8	2	1	100	790	46	64	12
9	2	1	101	795	2	23	84
10	2	2	104	775	28	7	31
11	2	2	105	780	84	21	66
12	2	2	106	785	10	73	47
13	3	1	104	800	65	83	39
14	3	1	105	795	75	59	65
15	3	1	106	790	96	69	17
16	3	2	109	770	45	53	91
17	3	2	110	775	47	48	29
18	3	2	111	780	48	42	41

Appendix B
SPSS Syntax File for the Illustrative Example

```
SET BLANKS=SYSMIS UNDEFINED=WARN printback=listing .
DATA LIST
  FILE='c:\spsswin\can_14.dta' FIXED RECORDS=1 /1
  id 1-2 exp_grp 4 gender 6 iq 8-10 gre 12-14
  x1 16-17 x2 19-20 x3 22-23 .
title 'show CCA is General Linear Model !!! can_14.sps !!!'.
execute .
list variables=all/cases=99999 .

subtitle 'compute orthogonal contrasts #################' .
execute .
compute a1 = 0 .
if (exp_grp eq 1) a1 = -1 .
if (exp_grp eq 2) a1 = 1 .
compute a2 = 0 .
if (a1 ne 0) a2 = -1 .
if (exp_grp eq 3) a2 = 2 .
if (gender eq 1) b1 = -1 .
if (gender eq 2) b1 = 1 .
compute a1_b1 = a1 * b1 .
compute a2_b1 = a2 * b1 .
print formats a1 TO a2_b1 (F3) .
list variabes=all/cases=99999 .
subtitle 'show contrasts orthogonal %%%%%%%%%%%%%%%%%%' .
execute .
correlations variables = a1 to a2_b1 .

subtitle 'show CCA subsumes regression $$$$$$$$$$$$$$$$$'.
execute .
regression variables=iq x1 TO x3/dependent=iq/
  enter x1 TO x3 .
manova x1 x2 x3 with iq/
  print=signif(multiv eigen dimenr)/
  discrim=stan corr alpha(.999)/.

subtitle 'CCA subsumes factorial ANOVA ======='.
execute .
anova iq by exp_grp(1,3) gender(1,2)/statistics=all .
subtitle 'b1  CCA subsumes factorial multi-way ANOVA ||||||'.
execute .
manova a1 a2 b1 a1_b1 a2_b1 with iq/
  print=signif(multiv).
subtitle 'b2  CCA subsumes factorial multi-way ANOVA $$$$$$'.
execute .
manova       b1 a1_b1 a2_b1 with iq/
```

```
  print=signif(multiv).
subtitle 'b3  CCA subsumes factorial multi-way ANOVA @@@@@@'.
execute .
manova a1 a2    a1_b1 a2_b1 with iq/
  print=signif(multiv).
subtitle 'b4  CCA subsumes factorial multi-way ANOVA ######'.
execute .
manova a1 a2 b1             with iq/
  print=signif(multiv).

subtitle 'CCA subsumes factorial MANOVA @@@@@@@'.
execute .
manova iq gre by exp_grp(1,3) gender(1,2)/
  print=signif(multiv) .
subtitle 'b1  CCA subsumes factorial multi-way MANOVA ||||||'.
execute .
manova a1 a2 b1 a1_b1 a2_b1 with iq gre/
  print=signif(multiv).
subtitle 'b2  CCA subsumes factorial multi-way MANOVA $$$$$$'.
execute .
manova       b1 a1_b1 a2_b1 with iq gre/
  print=signif(multiv).
subtitle 'b3  CCA subsumes factorial multi-way MANOVA @@@@@@'.
execute .
manova a1 a2    a1_b1 a2_b1 with iq gre/
  print=signif(multiv).
subtitle 'b4  CCA subsumes factorial multi-way MANOVA ######'.
execute .
manova a1 a2 b1             with iq gre/
  print=signif(multiv).
```