Categorical Variables in Regression: Implementation and Interpretation
By Dr. Jon Starkweather, Research and Statistical Support consultant

Use of categorical variables in regression analysis is often avoided due to confusion concerning interpretation. The purpose of this paper was to review four strategies for dealing with categorical variables in regression and demonstrate their use and interpretation. The paper is arranged in hierarchical fashion, from least complex strategy to most complex strategy, with the exception of criterion coding which is not very complex. By coincidence, this arrangement of topics also appears to reflect frequency of use for these procedures, from most often used to least often used strategies. The four strategies discussed in the paper are dummy coding, effects coding, orthogonal coding, and criterion coding (also known as criterion scaling).

All four strategies are used to allow researchers the ability to enter categorical predictor variables into a multiple regression analysis. All four strategies necessitate the creation of one or more variables to reflect the categories of the predictor variable. All four strategies reveal identical R^2. However, the interpretation of regression coefficients and conclusions drawn from them differs across each strategy.

For the demonstration of each strategy, a fictional data set was created consisting of one continuous outcome variable (DV_Score), one categorical predictor variable (IV_Group), and 15 cases (see Table 1). The predictor variable contains three groups; experimental 1 (value label 1), experimental 2 (value label 2), and control (value label 3). The demonstration of these four strategies will assume a functional knowledge of basic regression concepts and the statistical software SPSS. SPSS syntax is given in the Appendix so that the reader may replicate the demonstrations in this paper. The discussion that follows will deal with each strategy independently demonstrating its use and appropriate interpretation. Conclusions will close out the paper by drawing some comparisons between the four strategies and illustrate the strengths and weaknesses of each.

<div align="center">Discussion</div>

Dummy Coding

The name dummy coding seems appropriate in the sense that it is the simplest method of coding a categorical variable (Pedhazur, 1997). Dummy coding is used when a researcher wants to compare other groups of the predictor variable with one specific group of the predictor variable. Often, the specific group is called the reference group or category. So, for the example data here, we will compare each of the two experimental groups to the control group; the control group is our reference category. This illustration is used here because it represents a common expression of dummy coding in social science research, the comparison of treatment to control, with the control group being the reference category. It is important to note that dummy coding can be used with two or more categories. Dummy coding in regression is analogous to simple independent *t*-testing or one-way Analysis of Variance (ANOVA) procedures in that dummy coding allows the researcher to explore mean differences by comparing the groups of the categorical variable. In order to do this with regression, we must separate our predictor variable groups in a manner that allows them to be entered into the regression.

To accomplish this, we would create two new 'dummy' variables in our data set, labeled dummy 1 and dummy 2 (see Table 2). Each of these new variables is used to represent the presence of membership in a category of the predictor. The dummy 1 variable is used to identify all the members of the experimental 1 group. The dummy 2 variable is used to identify all the members of the experimental 2 group. To represent membership in a group on each of the

dummy variables, each case would be coded as 1 if it is a member and all other cases coded as 0. When creating dummy variables, it is only necessary to create $k - 1$ dummy variables where k indicates the number of categories of the predictor variable. All of the categories can be represented without having one dummy variable for each category. This highlights the meaning of the reference category, because it is always coded as 0. Furthermore, if there were one dummy variable for each category, we would be violating the assumption of no perfect collinearity (Hardy, 1993; see also Kirk, 1995).

At this point, it is necessary to address the importance of choosing a reference category. In the example used here, the reference category is easy to choose and often this is the case. The control group represents a lack of treatment and therefore is easily identifiable as the reference category. The reference category should have some clear distinction. However, much research is done without a control group. In those instances, identification of the reference category is generally arbitrary, but Garson (2006) does offer some guidelines for choosing the reference category. First, using categories such as *miscellaneous* or *other* is not recommended because of the lack of specificity in those types of categorizations (Garson). Second, the reference category should not be a category with few cases, for obvious reasons related to sample size and error (Garson). Third, some researchers choose to use a middle category, because they believe it represents the best choice for comparison; rather than comparisons against the extremes (Garson).

Now that we have our dummy variables constructed, we can include them in the regression analysis and interpret the results. In SPSS, the predictor variable would not be entered into the regression and instead the dummy variables would take its place. The results indicate a significant model, $F(2, 12) = 57.17, p < .001$. The regression model was able to account for 91% of the variance. Table 3 provides unstandardized regression coefficients (B), intercept (constant), standardized regression coefficients (ß), R, R^2, and adj.R^2.

Now, because dummy variables were used to compare experimental 1 ($M = 5.00, SD = 3.16$) and experimental 2 ($M = 12.00, SD = 3.16$) to the control ($M = 26.00, SD = 3.16$), the intercept term is equal to the mean of the reference category (i.e. the control group). Each regression coefficient represents the amount of deviation of the group identified in the dummy variable from the mean of the reference category (Pedhazur, 1997). So, some simple mathematics allows us to see that the regression coefficient for dummy 1 (representing experimental 1) is $5 - 26 = -21$. Also, the regression coefficient for dummy 2 (representing experimental 2) is $12 - 26 = -14$. All of this results in the regression equation:

$$\hat{Y} = 26.00 + (-21 * \text{dummy 1}) + (-14 * \text{dummy 2}). \tag{1}$$

Furthermore, we can divide each regression coefficient by its standard error, which is equal to the regression sums of squares degrees of freedom (*df*), and arrive at a *t*-score which can then be used with the residual sums of squares *df* to determine significance. This is indicated in the SPSS output as the significance test for each regression coefficient. Testing for significance of the regression coefficients is identical to running *t*-tests comparing the group means, which is essentially what would have been done in post-hoc testing if we had conducted a one-way ANOVA on this data.

Using the regression equation listed above (Equation 1), you arrive at predicted values for each case that are equal to the mean of the category that includes the case as a member. However, this only occurs when you only have dummy variables for one predictor in the model.

As the researcher specifies more predictor variables (continuous or categorical) in the model, the clean consistency of the example above evaporates. But, the underlying method and interpretation of dummy coding categorical variables for regression remains. For this reason, dummy coding represents a simple approach for primarily quasi-experimental designs.

Effect Coding

Effect coding is used when the researcher wants to compare categories of the predictor variable to the entire sample, represented by the grand mean of the outcome variable. Effect coding can be used with three or more categories. It is very similar to dummy coding in that we must create k -1 new (effect) variables, where k represents the number of categories of our predictor. Also like dummy coding, with effect coding we must choose a reference category, for the example here the control group is again used. There are no guidelines for choosing the reference category in effect coding, because we are not comparing the other categories to it. However, in effect coding, the reference category will be coded as -1 in all effect variables. Like dummy coding, membership in a category is designated by the code of 1 on the effect variable representing that category and the other categories are coded as 0. The exception to that rule is the reference category which, as mentioned above; is always coded as -1.

To accomplish this, we will create two new variables (see Table 4). Effect 1 represents membership in the experimental 1 category and effect 2 represents membership in the experimental 2 category. All cases coded as -1 represent membership in the control category. In this sense, dummy and effect coding are identical, because only two variables are able to capture all the information contained in our categorical predictor.

Now that we have our effect variables constructed, we can include them in the regression analysis and interpret the results. Again, in SPSS, the predictor variable would not be entered into the regression and instead the effect variables would take its place. The results indicate a significant model, $F(2, 12) = 57.17, p < .001$. The regression model was able to account for 91% of the variance. These model summary statistics are identical to the ones for the dummy code example. However, there are differences in the coefficients and how they are interpreted. Table 5 provides unstandardized regression coefficients (B), intercept (constant), standardized regression coefficients (ß), R, R^2, and adj.R^2.

Now, because effect variables were used to compare experimental 1 ($M = 5.00, SD = 3.16$), experimental 2 ($M = 12.00, SD = 3.16$), and control ($M = 26.00, SD = 3.16$) to the grand mean ($GM = 14.33, SD = 9.50$), the intercept term is equal to the grand mean of the outcome variable. Each regression coefficient represents the amount of deviation of the group identified in the effect variable from the grand mean (Pedhazur, 1997). So, some simple mathematics allows us to see that the regression coefficient for effect 1 (representing experimental 1) is 5 - 14.33 = -9.33. Also, the regression coefficient for effect 2 (representing experimental 2) is 12 - 14.33 = -2.33. All of this results in the regression equation:

$$\hat{Y} = 14.33 + (-9.33 * \text{effect 1}) + (-2.33 * \text{effect 2}). \qquad (2)$$

But, it is important to distinguish what those regression coefficients represent. Effect coding is so called, because each regression coefficient is the deviation of a given category from the grand mean. Pedhazur (1997) specifies this more clearly as "the deviation of a given treatment mean from the grand mean is defined as its effect" (p. 363). Recall that the general linear model specifies that each individual's score on an outcome variable is a function of three

elements; the grand mean plus the treatment effect for a given variable (or category in this case), plus error.

Using the regression equation listed above (Equation 2), you arrive at predicted values for each case that are equal to the mean of the category that includes the case as a member. The predicted values are the same as those for the first regression equation (Equation 1). Therefore, we can calculate the error for each case by subtracting the predicted score from the actual score. Using the resulting error term, or residual as it is called in the context of regression; we can place each case in the general linear model and confirm the treatment effect as the same as the regression coefficients for a given category. Because of this isolation of treatment effect, effect coding is usually reserved for experimental designs.

At first, it appears we have lost the information for the reference category. However, the effect of the control category is defined the same way as the effect of the other two categories. It is the deviation of the identified group mean (control in this case) from the grand mean. So, using the control condition mean of 26 and subtracting the grand mean of 14.33, we get a treatment effect of 11.67. It is important to keep in mind that the reference category in effect coding for regression is never included in the regression equation, but is included in the general linear model.

Remember that linear regression and ANOVA procedures are members of the general linear model family of analyses. The link between linear regression and ANOVA is mentioned here, because the significant R^2 displayed above in our model summary indicates that there is a significant difference among the three category means and the grand mean. As with ANOVA, the omnibus F tells us there is a significant difference, but further evaluation is necessary to locate where the differences lie. SPSS addresses this with tests of significance for each regression coefficient; essentially the regression coefficient is divided by the standard error of that coefficient to arrive at a t-score as mentioned previously. However, as Pedhazur (1997) reports "these tests are generally not used in the present context, as the interest is not in whether a mean for a given treatment or category differs significantly from the grand mean but rather whether there are statistically significant differences among the treatment or category means" (p. 367).

Orthogonal Coding

Orthogonal coding is used when a researcher wants to explore specific theory driven hypotheses about group differences among categories of a predictor variable. Orthogonal is defined as having absolutely no relationship. "If two variables are orthogonal, knowing the value of one variable gives no clue as to the value of the other; the correlation between them is zero" (Tabachnick & Fidell, 2001, p. 8). Orthogonal coding allows the researcher to specify coding values for testing specific categories against one another or a combination of categories. Essentially, "comparisons are orthogonal when the sum of the products of the coefficients for their respective elements is zero" (Pedhazur, 1997, p. 376). Like the two strategies mentioned above, orthogonal coding necessitates the creation of $k - 1$ new (orthogonal) variables, where k indicates the number of categories of our predictor variable. However, we have much more flexibility when coding values for our two new variables (orthog1 & orthog2) than in the previous two strategies. That flexibility is available because "for $k \geq 3$ means, there are an infinite number of sets of orthogonal contrasts, but each set contains only $k - 1$ orthogonal contrasts" (Kirk, 1995, p. 117). For our example here, the first contrast compares the experimental 1 category to experimental 2 (orthog1) and the second contrast compares the combination of experimental 1 and 2 to the control (orthog2). It is important to note that these

two comparisons represent one set of possible comparisons among an infinite number of sets. However, we must choose our coding values in such a way as to stay within the limits of orthogonality.

In order to do this, we must first identify codes for each comparison that are mutually exclusive and represent the comparisons we are attempting to investigate. So, for the first comparison, we chose codes to represent the comparison of the experimental 1 category (coded as -1) to the experimental 2 category (coded as 1), while ignoring the control category (coded as 0). For the second comparison, we chose codes to represent the comparison of the experimental 1 and 2 categories (both coded as 1) to the control category (coded as -2). The resulting code values are orthogonal, $[(1) + (-1) + (0)] + [(1) + (1) + (-2)] = 0$. Table 6 displays the newly created orthogonal variables.

Now that we have our orthogonal variables constructed, we can include them in the regression analysis and interpret the results. Again, in SPSS, the predictor variable would not be entered into the regression and instead the orthogonal variables would take its place. The results indicate a significant model, $F(2, 12) = 57.17$, $p < .001$. The regression model was able to account for 91% of the variance. These model summary statistics are identical to the ones for the dummy code and effect code examples. However, there are differences in the coefficients and how they are interpreted. Table 7 provides unstandardized regression coefficients (B), intercept (constant), standardized regression coefficients (ß), R, R^2, and adj.R^2. The resulting regression equation is:

$$\hat{Y} = 14.33 + (-3.50 * orthog1) + (-5.83 * orthog2). \tag{3}$$

Equation 3 yields predicted values for each case that are equal to the mean of the category that includes the case as a member. Again these predicted values are the same in the two previous examples.

Now, because we used orthogonal code values, the means of both orthogonal variables are zero, which in turn ensures that the intercept term is equal to the grand mean of the outcome variable ($GM = 14.33$). In this regard, effect coding and orthogonal coding are the same, but that is where the similarity stops. SPSS produces regression coefficients, but they are not of particular interest as is shown below. The regression coefficient for each orthogonal variable is calculated by dividing the sum of the cross products for the orthogonal variable and the outcome variable by the sum of the squared orthogonal variable values. However, the real value of orthogonal coding is that "the statistical tests of the regression coefficients can be interpreted as supporting or not supporting ones research hypotheses" (Henderson, 1998, p. 2). The research hypotheses referred to in the previous quote are the hypotheses that dictated which orthogonal comparisons were made. Therefore, according to our results, it appears that there was a significant difference between the two experimental categories, $t(12) = -3.50$, $p = .004$. Also, there was a significant difference between the combination of the two experimental categories and the control category, $t(12) = -10.10$, $p < .001$. Recall that the t-test for each regression coefficient is based upon the t-score provided when dividing the regression coefficient by its standard error.

One other interesting point is that because the orthogonal variable means' are zero, the standardized coefficients are the correlations of each orthogonal variable and the outcome variable. When these standardized coefficients are squared and summed the result is R^2, as would be expected in this example because; there are no other variables in the model.

Criterion Coding

Criterion coding is unique among the four strategies demonstrated here for a variety of reasons. First, when criterion coding, it is only necessary to create one new coded variable; criterion 1 for our example. The coded value for each case on the criterion variable is the mean of the category that includes the case as a member. Criterion coding was originally developed by Beaton (as cited in Schumacker & Williams, 1993), primarily as a way of decreasing the associated problems with large numbers of categorical predictors and/or categories among them which necessitated large numbers of coding variables. Schumacker and Williams also mention that other coding methods often necessitate exclusion of cases where data is missing from the categorical predictor, which is not necessary with criterion coding. Finally, Pedhazur (1997) states that criterion coding can also be utilized with ordinal variables.

Returning to our example, recall the means of each category from above; experimental 1 ($M = 5.00$, $SD = 3.16$), experimental 2 ($M = 12.00$, $SD = 3.16$), and control ($M = 26.00$, $SD = 3.16$). Using the mean of each category, we create the criterion variable (criterion 1; see Table 8).

Now that we have our criterion variable constructed, we can include it in the regression analysis and interpret the results. Again, in SPSS, the predictor variable would not be entered into the regression and instead the criterion variable would take its place. Notice that we are now dealing with bivariate regression, this will have implications below. The results indicate the regression model was able to account for 91% of the variance. The values for R, and R^2 are identical to the ones for all three examples above. However, there are differences in the other summary statistics, as well as the coefficients and how they are interpreted. Table 9 provides unstandardized regression coefficients (B), intercept (constant), standardized regression coefficients (ß), R, R^2, and adj.R^2.

First, the regression summary table in SPSS is incorrect. The table indicates that the model generated is significant, $F(1, 13) = 123.86$, $p < .001$. However, the df are incorrect which substantially increases the F value. Note, because criterion coding was used, the software only accounts for a single df for the regression sum of squares and accounts for 13 df for the residual sum of squares. Luckily the software reports the correct values for all the sum of squares. So, the researcher can correctly calculate the F ratio by using the correct df, recall that we are still utilizing three categories of the predictor and 15 cases. Therefore, the correct df are 2 and 12 for regression and residual, respectively. Utilizing the correct df, the resulting model is *still* significant, $F(2, 12) = 57.17$, $p < .001$. Notice that once corrected, the result is the same as all three mentioned above.

Next, notice that because the criterion coded variable was used and it consisted of the means for each category; the intercept value is 0. This will always be the case with criterion coding, because the regression coefficient is 1.0 and the grand mean ($GM = 14.33$) is equal to the mean of our newly created criterion coded variable. This too will always be the case, because the regression coefficient is produced by dividing the sum of the cross products of criterion variable values (actual dependent variable values) and the predicted values by the sum of the squared predicted values ($\sum y\hat{y}$)/($\sum \hat{y}^2$). In criterion coding, the sum of the cross products mention above and the sum of the squared predicted values are the same. This in turn is because the equations associated with each value will contain identical values for all the essential elements; namely the sum of predicted values squared ($\sum \hat{y}$)^2 will be the same as multiplying the sum of actual dependent variable values by the sum of predicted values ($\sum y$)($\sum \hat{y}$). The resulting regression equation is:

$$\hat{Y} = 0.00 + (1.00 * \text{criterion 1}) \tag{4}$$

Not surprisingly, Equation 4 yields predicted values for each case that are equal to the mean of the category that includes the case as a member, which are the same predicted values as those produced in all three examples above.

<div align="center">Conclusions</div>

      Four strategies were demonstrated as methods for coding a categorical predictor variable for inclusion in linear regression. Each offers specific utility for researchers implementing quasi-experimental designs and *true* experimental designs. It is important to note that each of these strategies resulted in identical values for model summary statistics. This would not be the case if multiple predictor variables were included in the model. Each of these strategies is compatible with multiple predictors, either continuous or categorical, which highlights the importance of understanding the differences associated with each strategy. Namely, the interpretation of regression coefficients differs across each strategy. It must also be noted that common regression practices using software dictate having the software compute (or save, as is the case with SPSS) the predicted values and error terms (standardized and unstandardized) for each case. The syntax included in the Appendix has the necessary elements to do this, and if used; will result in identical values for each case across the four strategies. This leads to a cautionary note about the use of categorical variables in regression. Given the preceding comment about the predicted values being the same across strategies, it should be clear that regression works best with continuous rather than categorical variables. However, if multiple predictors are included in the model, the use of categorical predictors becomes more *precise*. Because, instead of predicting the mean of each category (which was represented here due to only having one predictor), the predicted values resulting from the model will be based on all the variables included in the model.

Until next time; *won't you buy me a Mercedes Benz…?*

<div align="center">References</div>

Garson, G. D. (2006). *Statnotes: Topics in multivariate analysis: Multiple regression.* Retrieved June 3, 2010, from North Carolina State University, Department of Public Administration Web site: http://www2.chass.ncsu.edu/garson/pa765/regress.htm

Hardy, M. A. (1993). *Regression with dummy variables.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07-093. Newbury Park, CA: Sage.

Henderson, C. (1998). *The SPSS GLM procedure, or what happened to ANOVA?* Retrieved June 3, 2010, from University of North Texas, Computing and Information Technology Center, Benchmarks (July 1998) Web site: http://www.unt.edu/benchmarks/ archives/1998/july98/GLM.htm

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Harcourt Brace.

Schumacker, R. E. & Williams, J. D. (1993). Teaching ordinal and criterion scaling in multiple regression. *Multiple Linear Regression Viewpoints, 20,* 25 – 31.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needleham Heights, MA: Allyn and Bacon.

Table 1
Initial Example Data

Case	DV_Score	IV_Group
1	1	1
2	3	1
3	5	1
4	7	1
5	9	1
6	8	2
7	10	2
8	12	2
9	14	2
10	16	2
11	22	3
12	24	3
13	26	3
14	28	3
15	30	3

Table 2
Dummy Coded Example Data

Case	DV_Score	IV_Group	dummy1	dummy2
1	1	1	1	0
2	3	1	1	0
3	5	1	1	0
4	7	1	1	0
5	9	1	1	0
6	8	2	0	1
7	10	2	0	1
8	12	2	0	1
9	14	2	0	1
10	16	2	0	1
11	22	3	0	0
12	24	3	0	0
13	26	3	0	0
14	28	3	0	0
15	30	3	0	0

Table 3
Summary of Regression Analysis with Dummy Coded Variables
(N = 15)

Variable	B	SE B	β
Constant	26.00	1.41	
Dummy1	-21.00	2.00	-1.08*
Dummy2	-14.00	2.00	-0.72*

Note. $R = .95$, $R^2 = .91$, adj.$R^2 = .89$.
*p < .05.

Table 4
Effect Coded Example Data

Case	DV_Score	IV_Group	effect1	effect2
1	1	1	1	0
2	3	1	1	0
3	5	1	1	0
4	7	1	1	0
5	9	1	1	0
6	8	2	0	1
7	10	2	0	1
8	12	2	0	1
9	14	2	0	1
10	16	2	0	1
11	22	3	-1	-1
12	24	3	-1	-1
13	26	3	-1	-1
14	28	3	-1	-1
15	30	3	-1	-1

Table 5
Summary of Regression Analysis with Effect Coded Variables
(N = 15)

Variable	B	SE B	β
Constant	14.33	0.82	
Effect1	-9.33	1.16	-.83*
Effect2	-2.33	1.16	-.21

Note. $R = .95$, $R^2 = .91$, adj.$R^2 = .89$.
*p < .05.

Table 6
Orthogonal Coded Example Data

Case	DV_Score	IV_Group	orthog1	orthog2
1	1	1	1	1
2	3	1	1	1
3	5	1	1	1
4	7	1	1	1
5	9	1	1	1
6	8	2	-1	1
7	10	2	-1	1
8	12	2	-1	1
9	14	2	-1	1
10	16	2	-1	1
11	22	3	0	-2
12	24	3	0	-2
13	26	3	0	-2
14	28	3	0	-2
15	30	3	0	-2

Table 7
Summary of Regression Analysis with Orthogonal Coded Variables
(N = 15)

Variable	B	$SE\ B$	β
Constant	14.33	0.82	
Orthog1	-3.50	1.00	-.31*
Orthog2	-5.83	0.58	-.90*

Note. $R = .95$, $R^2 = .91$, adj.$R^2 = .89$.
*$p < .05$.

Table 8

Criterion Coded Example Data

Case	DV_Score	IV_Group	criterion1
1	1	1	5
2	3	1	5
3	5	1	5
4	7	1	5
5	9	1	5
6	8	2	12
7	10	2	12
8	12	2	12
9	14	2	12
10	16	2	12
11	22	3	26
12	24	3	26
13	26	3	26
14	28	3	26
15	30	3	26

Table 9

Summary of Regression Analysis with Criterion Coded Variable

(N = 15)

Variable	*B*	*SE B*	β
Constant	0.00	1.51	
Criterion1	1.00	0.09	-0.72*

Note. R = .95, *R²* = .91, adj.*R²* = .89.

*p < .05.

Appendix

SPSS Syntax

```
ONEWAY
  DV_Score BY IV_Group
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /MISSING ANALYSIS
  /POSTHOC = SCHEFFE ALPHA(.05).
comment: Dummy Coding Example.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT DV_Score
  /METHOD=ENTER Dummy1 Dummy2
  /SAVE PRED ZPRED RESID ZRESID .
comment: Effect Coding Example.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT DV_Score
  /METHOD=ENTER Effect1 Effect2
  /SAVE PRED ZPRED RESID ZRESID .
comment: Orthogonal Coding Example.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT DV_Score
  /METHOD=ENTER Orthog1 Orthog2
  /SAVE PRED ZPRED RESID ZRESID .
comment: Criterion Coding Example.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT DV_Score
  /METHOD=ENTER Criterion1
  /SAVE PRED ZPRED RESID ZRESID .
```