

# Performing Analyses of Variance with Continuous and Categorical Factors: The Easy Way

R. C. Gardner  
Department of Psychology  
University of Western Ontario

Often researchers are in a position where they have a set of categorical and continuous factors, and they wish to perform an analysis of variance. They have been advised that it is not appropriate for them to dichotomize, or otherwise form groups of the continuous variables, and that they must use multiple regression. This is certainly a viable solution, but it is not the only one. If the factors are all based on between subject variation (i.e., completely randomized designs), researchers can make use of the familiar SPSS GLM Univariate program and obtain results that are identical to those obtained with multiple regression, complete with regression coefficients, if desired.<sup>1</sup>

When using multiple regression, researchers can choose between:

1. ways of coding the categorical factors (often, but not always the distinction is made in terms of dummy or effect coding),
2. centering or not centering the continuous variables, and
3. considering all the factors and interactions at once (sometimes referred to as the unique solution, SSTYPE(3), or model 1 in traditional analysis of variance parlance (Overall & Spiegel, 1969), or using an hierarchical approach. There are variations on this last approach but a common one is where “main effects” are entered in a first set, “2-way interactions” on a second set, “3-way interactions” on a third set, etc. This is referred to as SSTYPE(2), or model 2.

We shall discuss here, a simple way to proceed that makes use of SPSS GLM Univariate. In the example to be investigated, there are three factors, a categorical factor (Grp with 3 levels), and two continuous variables that have been centered (cx, and cy). The dependent variable is identified as Z. A small sample of data is presented in Appendix A for those who may wish to copy the results presented here or to test the validity of the statement that the results will be exactly the same (with perhaps a few rounding errors) if they were to opt to do the analyses

---

<sup>1</sup> If they have a split plot factorial or repeated measures design, they can use the SPSS GLM Repeated Measures program using essentially the same procedures as described here.

using multiple regression. It should be noted, however, that the procedures described here can handle anything up to 5 factors (and neither of them has to be categorical necessarily).

We are not going to use multiple regression in any of these analyses, so the type of coding is not an issue. It should be noted, however, that if one were to use multiple regression, the type of coding is an issue. If one codes the data using effect coding and performs a unique solution (model 1), the F-ratios are correct but the interpretation of the regression coefficients is somewhat limiting. If one uses dummy coding for the same analysis, the F-ratios for the lower order effects are not correct, even though the regression coefficients are easier to interpret (what ever value there might be in such interpretations if the F-ratios are incorrect). In short, if one is using multiple regression for model 1, they should not use dummy coding if they want the correct answers. If one is performing model 2, the type of coding makes no difference, but the differences in interpreting the regression coefficients still remains.

#### Running SPSS GLM Univariate for Model 1

This is by far the easiest way to analyze the data. One simply specifies the dependent variable, identifies the categorical factor(s) as fixed factor(s) and identifies the continuous variables as covariates. It is then necessary to specify the model. This can be done, either by pasting the run into the syntax editor and adding all the possible interactions to the design statement, or by clicking on the *Model Window*, and then using the custom approach to build the whole model. If the regression coefficients for dummy coding are desired, one needs only to check *Parameter Estimates* in the *Options Window*. The Syntax file will look as follows.

```
UNIANOVA
z BY Grp WITH cx cy
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = ETASQ OPOWER PARAMETER
/CRITERIA = ALPHA(.05)
/DESIGN = Grp cx cy Grp*cx Grp*cy cx*cy Grp*cx*cy .
```

Running this produces the following analysis of variance summary table. It will be noted that this table presents not only the sums of squares and F-ratios, etc for the full model, but also the significance levels, partial eta squared values, and power estimates as for any analysis of variance table. This information can be obtained, of course, from multiple regression runs but it is much more laborious. The analysis of variance summary table is presented in Table 1.

Table 1.

## Tests of Between-Subjects Effects

Dependent Variable: z

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Corrected Model	661.946 <sup>b</sup>	11	60.177	4.499	.006	.792	49.486	.955
Intercept	4248.256	1	4248.256	317.591	.000	.961	317.591	1.000
Grp	41.088	2	20.544	1.536	.252	.191	3.072	.268
cx	23.916	1	23.916	1.788	.204	.121	1.788	.236
cy	48.863	1	48.863	3.653	.078	.219	3.653	.425
Grp * cx	81.004	2	40.502	3.028	.083	.318	6.056	.486
Grp * cy	40.659	2	20.329	1.520	.255	.190	3.040	.266
cx * cy	67.740	1	67.740	5.064	.042	.280	5.064	.549
Grp * cx * cy	2.896	2	1.448	.108	.898	.016	.216	.063
Error	173.894	13	13.376					
Total	21977.000	25						
Corrected Total	835.840	24						

a. Computed using alpha = .05

b. R Squared = .792 (Adjusted R Squared = .616)

As can be seen, the results of this analysis shows a significant ( $p < .042$ ) effect only for the two-way interaction between the two continuous factors (i.e.,  $cx * cy$ ). More on this approach later.

## Running SPSS GLM Univariate for Model 2

This analysis is performed in three stages. In the first stage, the analysis is run to assess the effects of the “main effects”. This is done by entering the dependent variable, the categorical variable(s) as fixed factor(s) and the centered variables as covariates and then running the analysis. This will produce the following Syntax file with no interaction terms which is, of course, simply the run for an analysis of covariance.

```
UNIANOVA
z BY Grp WITH cx cy
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = ETASQ OPOWER PARAMETER
/CRITERIA = ALPHA(.05)
/DESIGN = cx cy Grp .
```

Note, this is described as SSTYPE(3), but it is equivalent to SSTYPE(2) or model 2 simply because all the possible terms are not included in the analysis. The analysis of variance summary table for this run is presented in Table 2.

Table 2.

## Tests of Between-Subjects Effects

Dependent Variable: z

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Corrected Model	152.775 <sup>b</sup>	4	38.194	1.118	.376	.183	4.473	.287
Intercept	21083.280	1	21083.280	617.314	.000	.969	617.314	1.000
cx	44.115	1	44.115	1.292	.269	.061	1.292	.191
cy	126.718	1	126.718	3.710	.068	.156	3.710	.450
Grp	27.159	2	13.580	.398	.677	.038	.795	.105
Error	683.065	20	34.153					
Total	21977.000	25						
Corrected Total	835.840	24						

a. Computed using alpha = .05

b. R Squared = .183 (Adjusted R Squared = .019)

It will be noted that (a) the Sums of squares, mean squares and F-ratios are different than those reported for Model 1 (as they should be), and (b) none of the main effects are significant (nor were they in Table 1).<sup>2</sup>

The second stage of Model 2 involves adding the two-way interactions to the analysis, and adapting the Design statement as required. This will produce the following Syntax file.

```
UNIANOVA
z BY Grp WITH cx cy
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = ETASQ OPOWER PARAMETER
/CRITERIA = ALPHA(.05)
/DESIGN = Grp cx cy Grp*cx Grp*cy cx*cy .
```

Again, note that this is described as SSTYPE(3), but if we pay attention only to the sums of squares, etc., for the two-way interaction effects, they are in fact the estimates from model 2 (SSTYPE(2)) because the three-way interaction is not contained in the model. The analysis of variance summary table is presented in Table 3.

<sup>2</sup> Note that the error term for this analysis has 20 degrees of freedom and is the residual once the main effects have been estimated. Cohen, Cohen, West, & Aiken (2003, page 171) refer to this as a model 1 error term. If the error term for the full analysis were used (the one from Table 1 with 13 degrees of freedom), they would refer to it as a model 2 error term (somewhat of a reversal of terms in the present context).

Table 3.

## Tests of Between-Subjects Effects

Dependent Variable: z

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Corrected Model	659.050 <sup>b</sup>	9	73.228	6.213	.001	.788	55.918	.993
Intercept	5352.276	1	5352.276	454.120	.000	.968	454.120	1.000
Grp	50.337	2	25.169	2.135	.153	.222	4.271	.369
cx	32.727	1	32.727	2.777	.116	.156	2.777	.345
cy	276.675	1	276.675	23.475	.000	.610	23.475	.995
Grp * cx	82.666	2	41.333	3.507	.056	.319	7.014	.563
Grp * cy	158.753	2	79.376	6.735	.008	.473	13.470	.851
cx * cy	196.276	1	196.276	16.653	.001	.526	16.653	.968
Error	176.790	15	11.786					
Total	21977.000	25						
Corrected Total	835.840	24						

a. Computed using alpha = .05

b. R Squared = .788 (Adjusted R Squared = .662)

Because we are interested in Model 2 values, only the sums of squares, etc for the two-way interactions are of interest. In this case, note that the F-ratios for the Group\*cy and cx\*cy interactions are significant. Note too, that the so-called “main effect” for cy is also significant now, but this should be ignored because as we know when terms are added to a regression equation, the meaning and the values of the effects change.<sup>3</sup> More on interpretation, etc., later.

The third stage of model 2 involves adding the three-way interactions and modifying the Design statement to include all main and interaction effects. Because there are only three factors in this example, the results of this run will be identical to those for the model 1 results shown earlier. The Syntax file would be identical to that shown in that section.

Interpreting the Regression Coefficients and the *F*-ratios for Model 1

Although it is not meaningful to use dummy coding for factorial designs for model 1 when using multiple regression, the regression coefficients yielded by SPSS GLM Univariate in fact are those that one would obtain if dummy coding were used and the last group for each categorical factor was coded with all zeros. This apparent anomaly is because SPSS GLM Univariate does not actually perform a series of multiple regressions, but instead solves a set of matrix equations. Thus, the regression coefficients obtained from dummy coding are meaningful, and the F-ratios associated with the various main and interaction effects are correct.

<sup>3</sup> Again, note that the error term is the residual once all main effect and two-way interactions are estimated.

Calculating and Interpreting the Predicted values for each group.

The regression coefficients given for model 1 are presented in Table 4.

Table 4.

Parameter Estimates

Dependent Variable: z

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
					Lower Bound	Upper Bound			
Intercept	28.974	1.288	22.504	.000	26.192	31.755	.975	22.504	1.000
[Grp=1.00]	1.867	2.876	.649	.527	-4.345	8.079	.031	.649	.093
[Grp=2.00]	-5.869	3.878	-1.513	.154	-14.246	2.509	.150	1.513	.289
[Grp=3.00]	0 <sup>b</sup>	.	.	.	.	.	.	.	.
cx	-.753	.703	-1.072	.303	-2.271	.764	.081	1.072	.169
cy	.408	.525	.778	.450	-.725	1.542	.044	.778	.112
[Grp=1.00] * cx	1.081	1.478	.731	.477	-2.112	4.275	.040	.731	.104
[Grp=2.00] * cx	3.948	1.611	2.450	.029	.467	7.428	.316	2.450	.621
[Grp=3.00] * cx	0 <sup>b</sup>	.	.	.	.	.	.	.	.
[Grp=1.00] * cy	1.878	1.087	1.727	.108	-.471	4.227	.187	1.727	.360
[Grp=2.00] * cy	.877	1.852	.474	.644	-3.123	4.877	.017	.474	.072
[Grp=3.00] * cy	0 <sup>b</sup>	.	.	.	.	.	.	.	.
cx * cy	.754	.220	3.424	.005	.278	1.230	.474	3.424	.885
[Grp=1.00] * cx * cy	-.184	.451	-.407	.690	-1.157	.790	.013	.407	.067
[Grp=2.00] * cx * cy	.160	.914	.176	.863	-1.814	2.135	.002	.176	.053
[Grp=3.00] * cx * cy	0 <sup>b</sup>	.	.	.	.	.	.	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

The regression equation is:

$$Z = b_0 + b_1d1 + b_2d2 + b_3cx + b_4cy + b_5d1cx + b_6d2cx + b_7d1cy + b_8d2cy + b_9cxcy + b_{10}d1cxcy + b_{11}d2cxcy$$

where, the regression coefficients are as ordered in the table from  $b_0$  for the intercept to  $b_{11}$  for the regression coefficient for  $[Grp=2]*cx*cy$ . The equation can be reformulated to separate the estimate of the intercept from the slopes for  $cx$ ,  $cy$ , and  $cx*cy$  respectively. The values  $d1$  and  $d2$  represent the codes for the two vectors describing the three groups. The values for  $d1$  and  $d2$  are 1,0 for group 1, 0,1 for group 2, and 0,0 for group 3. The equation is:

$$\begin{aligned} Z = & (28.974 + 1.867d1 - 5.869d2) + \\ & (-.753 + 1.081d1 + 3.948d2)cx + \\ & (.408 + 1.878d1 + .877d2)cy + \\ & (.754 - .184d1 + .160d2)cx * cy \end{aligned}$$

which groups the terms together to compute the values for the intercept, slope for  $cx$ , slope for  $cy$ , and slope for  $cx*cy$  respectively. Solving the equation for each group yields the coefficients shown for each group in Table 5.

Table 5.

	Group 1	Group 2	Group 3
Intercept	30.841	23.105	28.974
Slope to cx	.328	3.195	-.753
Slope to cy	2.286	1.285	.406
Slope to cx*cy	.570	.914	.754

These values have considerable meaning. They are the coefficients that would be obtained if one were to perform multiple regression analyses for each group separately. They are also the values that would be obtained if multiple regression analysis were performed on all the data regardless of the coding used (dummy, effect, nonsense, etc...) for the categorical factor(s). And, finally for the coefficients referring to the slopes to cx and cy (but not to cx\*cy), they are the slopes when the other continuous variable is 0. That is, for group 1, the slope to cx of .328 is the value when cy = 0.

#### Interpreting the Regression Coefficients from Parameter Estimates.

Close inspection of these values in terms of the regression coefficients presented in the table of Parameter Estimates (Table 4) clarifies interpretation of the tests associated with those coefficients. Note, for example, that the intercept (28.974) in that table is the intercept for group 3, so the test of significance ( $t = 22.504$ ) indicates that the intercept for group 3 differs significantly from 0. Similarly, the regression coefficients for cx = -.753, cy = .408, and cx\*cy = .754 refer to the values for group 3, hence their tests of significance refer to deviations of these values from 0. Note, in the table of Parameter Estimates, that only the  $t$  for cx\*cy (3.424) is significant ( $p < .005$ ).

The interpretation of the coefficients involving the categorical factor is different because each one indicates deviations of the values for groups 1 and 2 from group 3. Thus, the regression coefficients for [Grp=1] and for [Grp = 2] are equal to  $1.867 = 30.841 - 28.974$ , and  $-5.869 = 23.105 - 28.974$  respectively so that the corresponding  $t$ -tests assess the significance of

these contrasts. Similar observations apply to the slopes for cx, for cy, and for cx\*cy.

Interpretation of the F-ratios from the analysis of variance.

The F-ratios can be interpreted in terms of these coefficients, as follows:

- Intercept: F = 317.591, p<.001: A test that the mean intercept (27.640) differs significantly from 0.
- Grp: F = 1.536, ns. A test that the intercepts (30.841, 23.105, and 28.974) differ more than would be expected by chance.
- cx: F = 1.788, ns. A test that the mean slope for cx (.923) differs significantly from 0.
- cy: F = 3.653, ns. A test that the mean slope for cy (1.326) differs significantly from 0.
- Grp\*cx F = 3.028, ns. A test that the slopes for cx (.328, 3.195, and -.753) differ more than would be expected by chance.
- Grp\*cy F = 1.520, ns. A test that the slopes for cy (2.286, 1.285, and .406) differ more than would be expected by chance.
- cx\*cy F = 5.064, p<.042 A test that the mean slope for cx\*cy (.746) differs significantly from 0. Because this is the only effect that was significant, only this interaction would be investigated. Note too that the t-value associated with this term refers only to the interaction in group 3.
- Grp\*cx\*cy F = .108, ns. A test that the slopes for cx\*cy (.570, .914, and .754) differ more than would be expected by chance.

Interpreting the Regression Coefficients and the F-ratios for Model 2

The F-ratios and regression coefficients are considered separately for each stage.

Stage 1. The regression coefficients are presented in Table 6.

Table 6.

Parameter Estimates

Dependent Variable: z

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
					Lower Bound	Upper Bound			
Intercept	28.837	2.019	14.282	.000	24.625	33.048	.911	14.282	1.000
cx	.890	.783	1.137	.269	-.744	2.524	.061	1.137	.191
cy	.879	.456	1.926	.068	-.073	1.831	.156	1.926	.450
[Grp=1.00]	2.080	2.919	.712	.484	-4.010	8.169	.025	.712	.104
[Grp=2.00]	-1.319	3.633	-.363	.720	-8.897	6.259	.007	.363	.064
[Grp=3.00]	0 <sup>b</sup>	.	.	.	.	.	.	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

The regression equation is:

$$Z = b_0 + b_1cx + b_2cy + b_3d1 + b_4d2$$



Rearranging terms shows that the intercepts are:  $b_0 + b_3d_1 + b_4d_2$ , while the slopes are  $b_1$  for  $c_x$  and  $b_2$  for  $c_y$ . Defining  $d_1$  and  $d_2$  as 1,0 for group 1, 0,1, for group 2, and 0,0 for group 3, yields the coefficients in Table 7.

Table 7

	Group 1	Group 2	Group 3
Intercept	30.917	27.518	28.837
slope for $c_x$	.890	.890	.890
slope for $c_y$	.879	.879	.879

Examination of these values demonstrates the meaning of the regression coefficients reported in Table 6, Parameter Estimates. Thus, for dummy coding, the intercept (28.837) is the intercept for group 3, and the regression coefficients for  $c_x$  and  $c_y$  are the slopes for group 3 (which don't differ in the other groups because at this point, the interaction is not part of the equation, and thus is set at 0). Furthermore, as noted with respect to the coefficients for model 1, the regression coefficients for [Grp = 1] (2.080) and [Grp = 2] (-1.319) are contrasts of the values for each group from group 3. That is,  $30.917 - 28.837 = 2.080$ , and  $27.518 - 28.837 = -1.319$ . The  $t$ -tests of these coefficients are therefore tests of these contrasts.

An added advantage of running the analysis through SPSS GLM Univariate is that at this stage, it is simply an analysis of covariance with two covariates. Thus, you can have the computer perform tests of the contrasts involving the three adjusted means (i.e., the intercepts in Table 7), and obtain information on all three contrasts, not just the two identified by the regression coefficients for groups in Table 6.

The  $F$ -ratios from Stage 1 of the analysis of variance for model 2 are tests of the estimates made at this stage. Their values and interpretation are as follows:

Intercept:	$F = 617.314, p < .001$ :	A test that the mean intercept (29.091) differs significantly from 0, assuming no interactions.
Grp:	$F = .398, ns$ .	A test that the intercepts (30.917, 27.518, & 28.837) differ more than expected by chance, assuming no interactions.
$c_x$ :	$F = 1.292, ns$ .	A test that the mean slope for $c_x$ (.890) differs significantly from 0, assuming no interactions.

cy:  $F = 3.710$ , ns. A test that the mean slope for cy (.879) differs significantly from 0, assuming no interactions.

Stage 2. This stage of model 2 adds the two way interactions to the equation, resulting in the regression coefficients presented in Table 8.

Table 8.

Parameter Estimates

Dependent Variable: z

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
					Lower Bound	Upper Bound			
Intercept	28.978	1.208	23.979	.000	26.402	31.554	.975	23.979	1.000
[Grp=1.00]	2.394	2.444	.979	.343	-2.817	7.604	.060	.979	.151
[Grp=2.00]	-5.543	3.368	-1.646	.121	-12.722	1.636	.153	1.646	.338
[Grp=3.00]	0 <sup>b</sup>	.	.	.	.	.	.	.	.
cx	-.697	.636	-1.097	.290	-2.052	.658	.074	1.097	.177
cy	.367	.475	.772	.452	-.646	1.380	.038	.772	.112
[Grp=1.00] * cx	1.246	1.346	.926	.369	-1.623	4.116	.054	.926	.140
[Grp=2.00] * cx	3.760	1.423	2.642	.018	.727	6.794	.318	2.642	.695
[Grp=3.00] * cx	0 <sup>b</sup>	.	.	.	.	.	.	.	.
[Grp=1.00] * cy	2.237	.623	3.589	.003	.908	3.565	.462	3.589	.918
[Grp=2.00] * cy	1.274	.969	1.315	.208	-.791	3.339	.103	1.315	.234
[Grp=3.00] * cy	0 <sup>b</sup>	.	.	.	.	.	.	.	.
cx * cy	.719	.176	4.081	.001	.344	1.095	.526	4.081	.968

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

The regression equation is:

$$Z = b_0 + b_1d1 + b_2d2 + b_3cx + b_4cy + b_5d1cx + b_6d2cx + b_7d1cy + b_8d2cy + b_9cx * cy$$

Rearranging the terms to identify the intercepts and slopes for the three groups yields Table 9.

Table 9.

	Group 1	Group 2	Group 3
Intercept	31.372	23.435	28.978
Slope to cx	.549	3.063	-.697
Slope to cy	3.604	1.641	.367
Slope to cx*cy	.719	.719	.719

Note, the intercepts are different from those reported in Table 7 (as they should be) and that the slopes to cx and cy vary across the 3 groups, while the slope to cx\*cy does not. The

interpretation of these values is similar to that for the other tables, though the values for the intercepts should not be considered because they represent values assuming possible two-way interactions but no three-way interaction. The interpretations for the other values are, however, as before. Similarly, the interpretation of the regression coefficients for the two-way interaction components in the previous table of Parameter Estimates is the same as before, but involves the values for each group in the above table. Note, furthermore, that the slope for  $cx*cy$  is constant over the three groups, because there is no three-way interaction term in the equation (i.e., the three-way interaction is assumed to be 0 at this point).

#### Interpreting the $F$ -ratios from the Analysis of Variance.

Only the  $F$ -ratios associated with the two-way interactions are interpretable at this stage.

They are as follows:

Grp*cx	$F = 3.507, ns$	A test that the slopes for cx (.549, 3.063, and -.697) differ more than would be expected by chance, assuming no three- way interaction.
Grp*cy	$F = 6.735, p<.008$	A test that the slopes for cy (3.604, 1.641, and .367) differ more than would be expected by chance, assuming no three- way interaction.
cx*cy	$F = 16.653, p<.001$	A test that the mean slope for cx*cy (.719) differs significantly from 0, assuming no three way interaction.

Stage 3. The analysis of variance and the regression coefficients at this stage are identical to those for model 1 in this example because there are only three factors. If the intent was to report model 2 results, then only the  $F$ -ratio for the three way interaction and the regression coefficients for  $[grp=1]*cx*cy$  and  $[grp=2]*cx*cy$  should be reported and interpreted.

#### Two Models: Two Sets of Results

The analysis for model 1 detected only a two-way interaction between cx and cy,  $F(1,13)=5.064, p<.042$ . The analysis for model 2 identified two significant effects. One was the  $cx*cy$  interaction,  $F(1,15)=16.653, p<.001$ , and the other was the Group\*cy interaction,  $F(2,15) = 6.375, p<.008$ . It shouldn't come as a surprise that the models yield different results, because they are in fact different models. For these data, model 1 indicates that the  $cx*cy$  interaction adds significantly to all the other effects in the design, while model 2 indicates that  $cx*cy$  and Group\*cy each add to the other effects, assuming there is no variation associated with the three-way interaction. Thus, there are important differences between models 1 and 2 that researchers

should be aware of and understand when performing these analyses. My preference is model 1 because it considers effects when all else is in the model, but there are different perspectives in this regard.

#### References

Cohen, J., Cohen, P., West, S.G. & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (Third edition)*. Mahwah, NJ: Lawrence Erlbaum.

Overall, J.E. & Spiegel, D.K. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin*, 72, 311-322.

## Appendix A

## Sample data

Grp	Z	cx	cy
1.00	31	-3.24	4.04
1.00	28	-3.24	-3.96
1.00	34	-2.24	3.04
1.00	33	-1.24	2.04
1.00	36	-1.24	3.04
1.00	27	-.24	-2.96
1.00	24	-.24	-1.96
1.00	24	-.24	-2.96
2.00	23	.76	-2.96
2.00	24	1.76	-.96
2.00	21	1.76	-1.96
2.00	35	1.76	2.04
2.00	36	1.76	2.04
2.00	19	2.76	-1.96
2.00	38	3.76	-.96
2.00	37	3.76	1.04
3.00	36	-2.24	-3.96
3.00	32	-1.24	-1.96
3.00	28	-.24	-.96
3.00	22	1.76	-2.96
3.00	28	.76	2.04
3.00	32	.76	1.04
3.00	24	-4.24	4.04
3.00	21	-3.24	3.04
3.00	34	1.76	3.04