Interactions between Binary & Quantitative Predictors

These data are from a treatment outcome study. Folks who were diagnosed as depressed were given the opportunity to begin treatment immediately, or to wait until they "feel that it would be a good time to begin therapy". Treatment was either a "peer support group" (code = 1) or "traditional cognitive-behavioral therapy" (code = 2). The researcher also recorded the time between when the diagnosis was made and when the assigned treatment was started, in days (delay). Treatment lasted 3 months, and then the outcome variable (dep -- a depression measure, so higher scores are "poorer") was assessed by a panel of clinicians who were "blind" as to the treatment each patient was receiving.

Here are the group means for Depression and Time.

There is a small group difference for Depression.

There is also a small group difference for Time

Report

treatment condition		rating of depression -- bigger scores are poorer	time between diagnosis and rating of depression - in days
support group	Mean	4.4000	32.6000
	N	20	20
	Std. Deviation	1.81804	17.43680
cog-beh group	Mean	4.0000	29.9000
	N	20	20
	Std. Deviation	2.00000	17.95286
Total	Mean	4.2000	31.2500
	N	40	40
	Std. Deviation	1.89737	17.52178

Dummy-Coded Regression of the Data

Here's the SPSS syntax code to dummy code the grouping variable, to center the covariate (using mean overall mean for that variables from above) and to compute the interaction term.

- Dummy coding follows the GLM convention – the group with the highest original code as the control group
- Centering of quantitative variables reduces "unnecessary" collinearity and simplifies interpretation of b
- Interactions are "non-additive combinations" -- meaning products of the related main effects


```
Syntax1 - SPSS Syntax Editor

File  Edit  View  Analyze  Graphs  Utilities  Run  Window
Help

if (grp = 1) grp_d = 1.
if (grp = 2) grp_d = 0.

compute cov_c = cov - 31.25.

compute int = grp_d * cov_c.
exe.
```

IF statements to dummy-code the group variable: PEER is coded "1" as the target group and CBT is coded "0" as the comparison group

Centering the covariate requires subtracting the mean from each person's COV score

The product of the coded group variable and the centered covariate is the interaction term

Below are several models (and some other stuff):
- An "ANCOVA" or "main effects" model with both the grouping variable and the covariate
- A short tirade about the limitations of the "main effects model" and the importance of interactions
- An "Interaction test" using nested model comparison
- A "full model ANCOVA" or "interaction model" with the grouping variable, the covariate and their interaction (and how to get all the information from the interaction test from a single model)
- Using GLM to obtain the full model

"Main Effects Model" or "ANCOVA"   via Regression

Model Summary

Model	R	R Square
1	.126[a]	.016

a. Predictors: (Constant), COV_C, GRP_D

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.219	2	1.109	.297	.745[a]
	Residual	138.181	37	3.735		
	Total	140.400	39			

a. Predictors: (Constant), COV_C, GRP_D

b. Dependent Variable: rating of depression -- bigger scores are poorer

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constan	3.990	.433		9.220	.000
	GRP_D	.419	.613	.112	.684	.498
	COV_C	-.007	.018	-.067	-.407	.686

a. Dependent Variable: rating of depression -- bigger scores are p

a is the expected dep when when the value of group and the covariate are both 0. Which is the mean for members of the comparison group (CBT - coded 0) with exactly the mean time (which has been centered to 0)

COV_C is the slope of the delay-DV regression line for each group (model assumes there is no interaction between GRP and COV) -- after correction for the IV

GRP_D is the direction and extent of the mean depression difference between the target group (PEER) and the comparison group (CBT) -- after correction for the covariate at its mean. The target group (PEER)mean is .419 larger than the comparison group (CBT).

Why run the ANCOVA or Multiple Regression model?

When we looked at the means of the DV and delay for the two groups, we noticed that there was a mean difference between the groups on delay. This means that the relationship between group membership the DV is "confounded" by the group difference on delay.

So, following the our general principle that when we don't have an experimental data "bigger models are more accurate portrayals of complex behavior" we included both group and delay in this model.

If we find a different group effect with and without the delay included, we expect that the model with it included is more likely to be accurate. If, as in this case, the group effect is the same with and without delay effect, we gain confidence in the generalizability of the group effect.

## A Short Tirade About Main Effects Multiple Regression Models

Until now, all the multiple regression models we've looked at have been "main effects" models – that is, we've included any variables we want, without including any interactions.

When interpret the Group regression weight from this ANCOVA or Multiple Regression model we are invoking the "homogeneity of regression slope" assumption, which is another way of saying this is that we are assuming there is no interaction between the groping variable and the covariate.

- In ANCOVA langage → the slope of the DV-Cov regression line is the same for all groups
- In ANOVA language → the Group difference on the Dv is the same for all values of delay

This is the major historical difference between the uses of ANOVA and Multiple Regression. If we have multiple (qualitative) predictors we put them in an ANOVA and the very first thing we would look at is the interaction. If there is an interaction, we would very carefully examine the main effects, to see if they are descriptive or misleading. However, if we run a multiple regression, whether we have quantitative, coded categorical or a mixture of variables types, we put them all into the model and blithely interpret their regression weights (main effects) without considering that there might be interactions and that the interpretation of the regression weights might be conditional upon the values of the other variables.

I've written a "computator" that computes plotting points to draw the regression lines from various models.
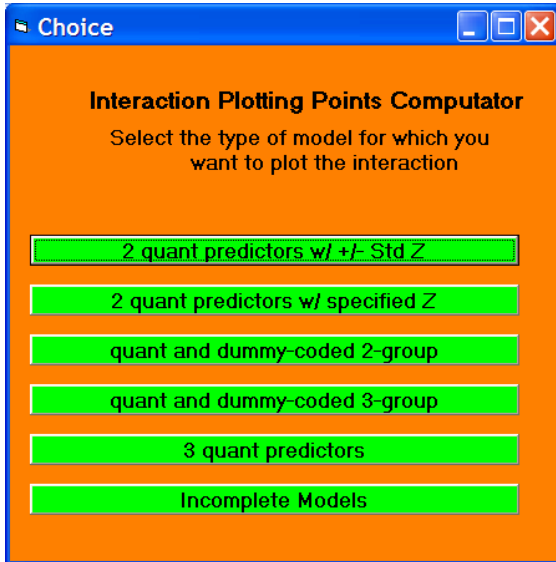
Select the one for 2 dummy-coded groups.

It was written for models including the interaction, but will work for the main effect & ANCOVA model also.
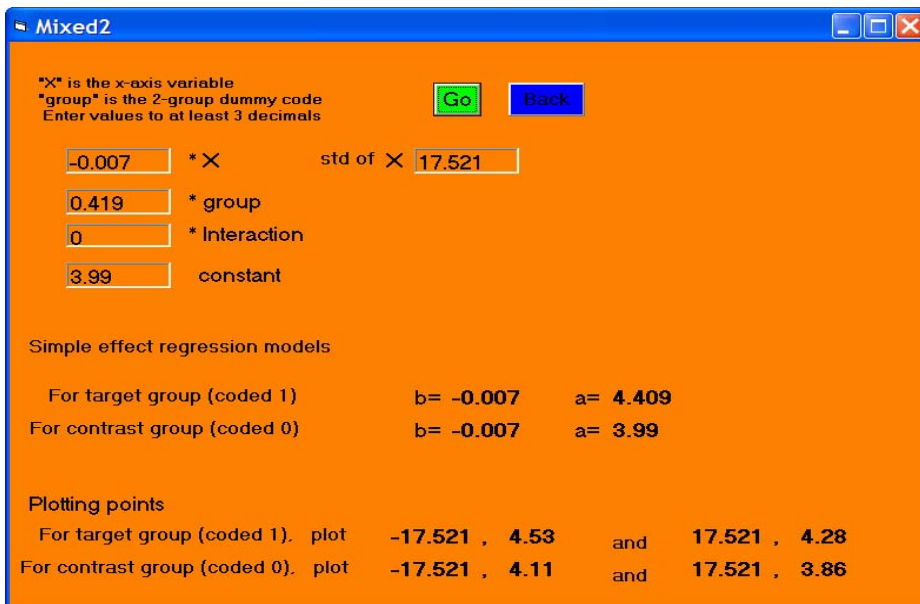
"X" is the covariate

the interaction weight will be "0"

get the standard deviation of the covariate from above

enter the info and click "Go"

The program gives you the regression models for each groups (notice they have the same slope) and the plotting points (based on +/- 1 std for the centered covariate)

Notice that the regression slope for each group's regression line is the same as the slope of COV from the full model (because we assume there's no interaction, don't put one in, and so have parallel regression lines for the two groups)

Notice that the difference between the group's constants is the same as the corrected mean difference. Notice also it is the same for all values of the covariate – because there's no interaction.
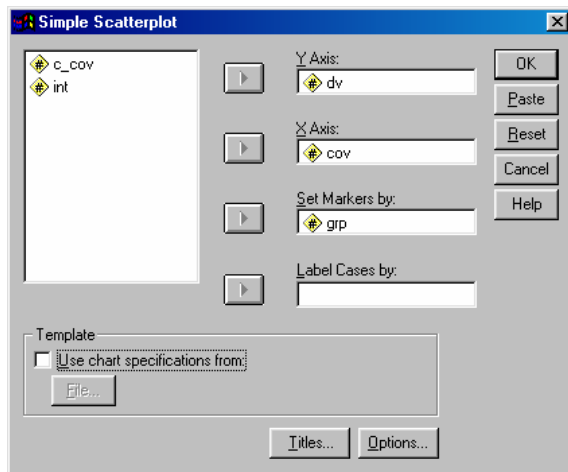
You should also remember that the lines are pretty flat and that the R² is < .02 -- this model isn't doing very well.
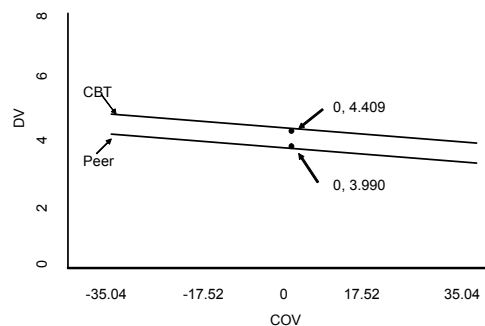
This graph is the "model" of the data. However, we know that it fits the data poorly ($R^2 < 2\%$). Maybe we should look at the data.
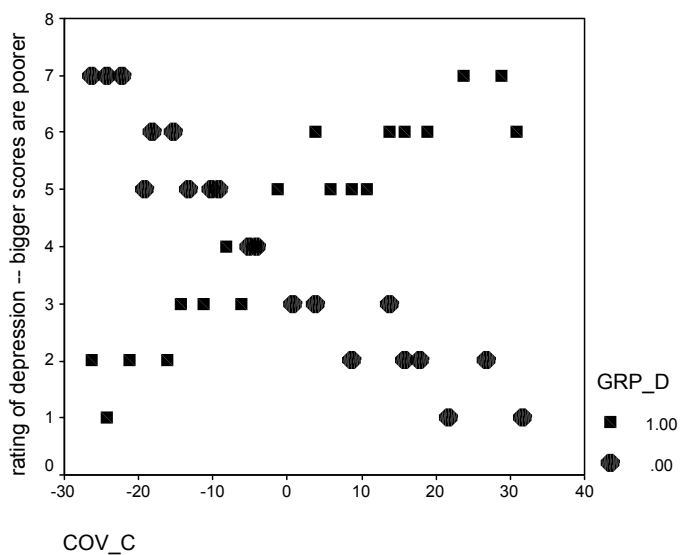
To graph the data…

Graph → Scatter → highlight the "Simple" scatterplot and click on Define

By convention, the DV goes on the Y axis and the covariate goes on the X axis.

Identify each plotted case with its group membership by setting the Marker to be "grp".

Add titles and footnotes by clicking the "Titles" button

This plot looks very different from the model above.

Here it looks like the regression lines for the two groups have very different slopes -- suggesting an interaction.

"Full Model" or "ANCOVA w/ Interaction" via Regression

Testing an interaction requires we include an "interaction term" in the model. All interactions are "non-additive combinations" of main effects. Non-additive means "multiplicative" -- that is we use product terms to represent interactions. The syntax for this is shown above (be sure to use the dummy-coded group variable and the centered covariate). The interpretation of the resulting regression weight (b) is important:

+   interaction regression weights indicate that the regression line of the target group (coded = 1) has a "more positive" slope than does the regression line of the comparison group (coded = 0). A more positive slope for the higher coded group.

-   interaction regression weights indicate that the regression line of the target group (coded = 1) has a "less positive" slope than does the regression line of the comparison group. (coded = 0). A more positive slope for the loser coded group.

nsig   interaction regression weights indicate that the regression lines of the two groups are parallel (no int). Same slope for the two groups.

Using Regression to Obtain an Interaction Model

There are a two common ways to perform the multiple regression with an interaction term:

- A full model with all the terms added
  - Since each "effect" in the model is a single term, the t-test of the associated b tells whether or not that term contributes to the model after correction for the other terms in the model
  - If you get the semi-partial correlations with the regression output you can compute the $R^2\Delta$ as the square of the semi-partial correlation of the interaction

- A 2-step hierarchical model with the main effects included first (COV and IV) and then the Interaction
  - Proponents of this often like having the $R^2\Delta$ to summarize the independent contribution of the interaction and the $R^2\Delta$F-test to test that contribution.

Full Regression Model Approach

Model Summary

Model	R	R Square
1	.952[a]	.907

a.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	127.332	3	42.444	116.93	.000[a]
	Residual	13.068	36	.363		
	Total	140.400	39			

a. Predictors: (Constant), INT, GRP_D, COV_C

b. Dependent Variable: rating of depression -- bigger scores are poorer

As expected, this looks just like the last step of the hierarchical modl.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations		
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part
1	(Constant)	3.856	.135		28.536	.000			
	grp_d	.411	.191	.110	2.153	.038	.107	.338	.109
	cov_c	-.107	.008	-.159	-13.871	.000	-.058	-.318	-.066
	int	.205	.011	.804	18.565	.000	.636	.952	.944

a. Dependent Variable: rating of depression -- bigger scores are poorer

Notice again that results of the full model are the same as the second step of the hierarchical approach below.

Also, the $R^2\Delta$ attributable to the interaction is computed as the square of the part correlation of the interaction

Looking at the regression weights from the full model:

Dummy code weights tell us about the height difference of the regression lines for the two groups at the mean of the covariate. The regression weight for group dummy code (CBT = 0, PEER = 1) tells us that the peer group has a mean depression score that is .411 larger than the CBT group, when holding time and the group*time interaction constant at zero (the mean of the centered variable). This mean difference is statistically significant.

Covariate weights tell us about the slope of the regression line for the control group. The regression weight for the centered covariate tells that depression is expected to decrease .107 for each additional day before the participant starts therapy, holding the other variables constant. Specifically, this is the slope of the regression line for those in the comparison group (PEER) with a score of "0" on the interaction term.

Interaction weights tell us about the difference in slope between the regression lines of the two groups. The interaction regression weight tells us that the slope of the depression-time regression line for the CBT group (1) is .205 greater than the slope of the depression-time regression line for the PEER group (0).

Things to Notice:

- The regression weight for group is actually a tiny bit smaller in the interaction model (.411) than in the main effects model, but is significant (p=.03) in the interaction model. Why? A version of multivariate power. Adding the interaction raised the R^2 from .016 to .907, and decreased the MSerror from 3.735 to .363. The regression weights in the interaction model are tested much more powerfully with this much smaller error term.
- The regression weight for delay is slightly larger (more negative) in the interaction model and is significant, for the same basic reason as for the group effect being significant in the interaction model.

There are a couple of things about which we must be very careful…

- The group differences is tested specifically at the mean of the quantitative variable (0 because of centering). We know nothing about the group difference at other values of the quantitative variable from the regression weight.
- The significant interaction tell us that the direction and/or size of the group difference *depends upon* the specific value of the quantitative variable.
- When there is a significant interaction, we must examine the main effect very carefully, to determine if it is "descriptive" (unconditionally true for all values of the quantitative variable) or "potentially misleading" (conditionally true for only some values of the quantitative variable)

The "Data Picture"

It is important to get a sense for the statistical model represented by the regression weights. That is why it is important to plot the model, and be able to describe the "story" that it tells.
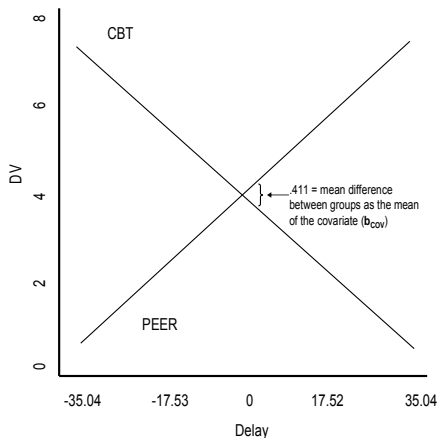

Again, select "quant and dummy-coded 2-group"

"X" is the covariate

Enter the standard deviation of the covariate (from above)

Enter the info and click "Go"

The program gives you the separate regression model for each group and the plotting points (based on +/- 1 std for the centered covariate)

This model is a much better approximation of the plot of the raw data shown above.

The LSD$_{mmd}$ tells us the minimum mean group difference that is significant.

Using it we can identify a "zone of significance" showing for what values of Delay there are significant treatment group differences

Looking at t[...] grate and make sense of the regression weights.

There is an interaction of treatment group and delay
- there is a negative relationship between delay and depression for those in the CBT group
  o those who delayed longer had higher depression scores than those who delayed a shorter time
- whereas there was a positive relationship between delay and depression for those in the PEER group
  o those who delayed longer had lower depression scores than those who delayed a shorter time

Overall there is a positive relationship between delay and depression
- those who delayed longer had lower depression scores than those who delayed a shorter time
- however this is descriptive only for those in the PEER group, and not for the CBT group

The mean of the target group (PEER) is significantly higher than the mean of the comparison group (CBT), when holding the covariate constant at its mean
- based on the LSD$_{mmd}$ this difference is descriptive only for values of Delay from about 31 and above
- there is no depression difference between the groups for values of delay from about 28.5 to about 31
- the CBT group has a significantly higher depression score for delay values less than about 28.5

Hierarchical Regression Approach

This is the "classic" approach and still preferred by many. SPSS Regression was used to obtain first the main effects model and then add the interaction term to obtain the full model. $R^2\Delta$ tests and semi-partial correlations were also requested.

Model Summary

Model	R	R Square	R Square Change	F Change	df1	df2	Sig. F Change
				Change Statistics			
1	.126[a]	.016	.016	.297	2	37	.745
2	.952[b]	.907	.891	344.677	1	36	.000

a. Predictors: (Constant), COV_C, GRP_D

b. Predictors: (Constant), COV_C, GRP_D, INT

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.219	2	1.109	.297	.745[a]
	Residual	138.181	37	3.735		
	Total	140.400	39			
2	Regression	127.332	3	42.444	116.930	.000[b]
	Residual	13.068	36	.363		
	Total	140.400	39			

a. Predictors: (Constant), COV_C, GRP_D

b. Predictors: (Constant), COV_C, GRP_D, INT

c. Dependent Variable: rating of depression -- bigger scores are poorer

As we saw above, the main effects ANCOVA model didn't account for much variance and neither the COV nor the grouping variable contributed to that model.

Adding the interaction term increased the R^2 considerably.

Also with this large increase in the fit of the model, the error term decreased dramatically, so that the main effects of the COV and GRP are both significant.

Notice that squaring part correlation of the interaction in the full model ($.944^2 = .891$) gives the $R^2\Delta$ from adding the interaction to the main effect model.

The second step shows exactly the same results as the full model above.

Coefficients[a]

| Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Partial	Part
1	(Constant)	3.990	.433		9.220	.000			
	grp_d	.419	.613	.112	.684	.498	.107	.112	.112
	cov_c	-.007	.018	-.067	-.407	.686	-.058	-.067	-.066
2	(Constant)	3.856	.135		28.536	.000			
	grp_d	.411	.191	.110	2.153	.038	.107	.338	.109
	cov_c	-.107	.008	-.159	-13.871	.000	-.058	-.318	-.066
	int	.205	.011	.804	18.565	.000	.636	.952	.944

a. Dependent Variable: rating of depression -- bigger scores are poorer

Interaction Models using SPSS GLM

There are different ways to coerce GLM into computing this model. The following variation gives all the information provided by the multiple regression approach plus giving the group means after correction for the covariate and the interaction.

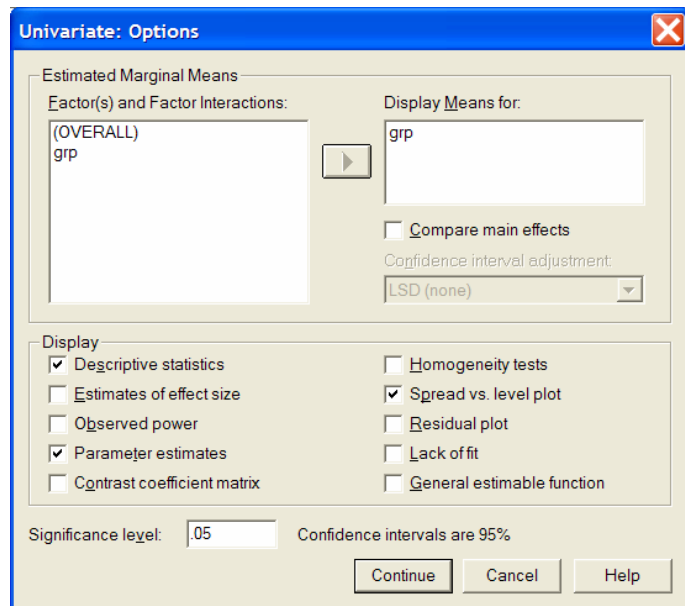**Analyze ➔ General Linear Model ➔ Univariate**
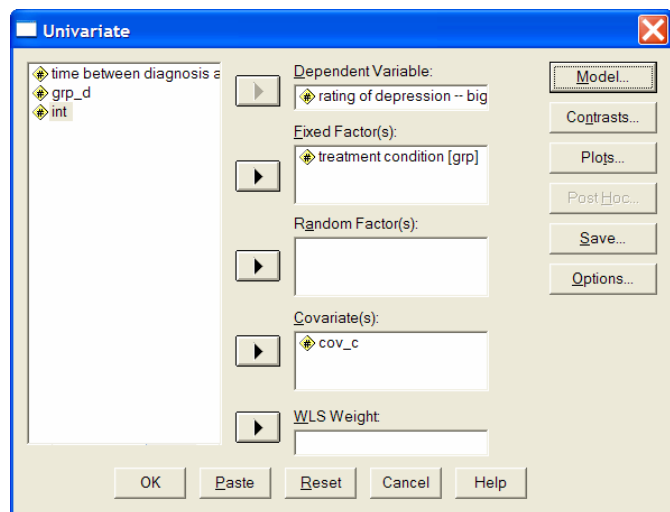
Use the original group variable as the Fixed Factor

Be sure to use the centered delay term as the "Covariate"
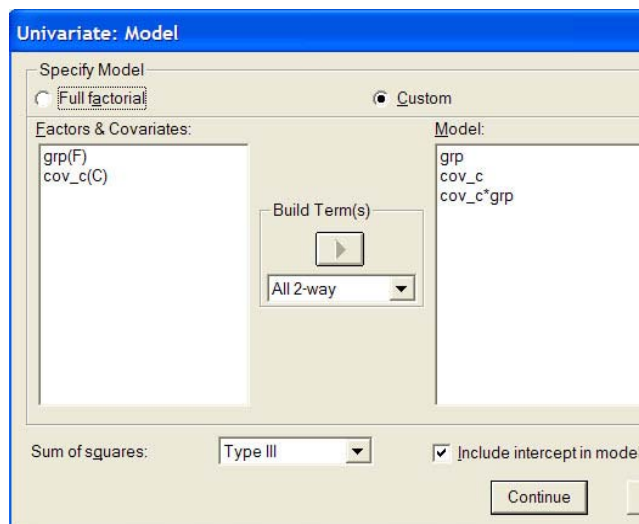
Put the grouping variable and the covariate in as main effects

Combine them into a 2-way interaction

Move the IV into the "Display Means for" window -- to obtain the corrected group means

Be sure to check Descriptive Statistics and Parameter estimates

Descriptive Statistics

Dependent Variable: rating of depression -- bigger scores are poorer

treatment condition	Mean	Std. Deviation	N
support group	4.4000	1.81804	20
cog-beh group	4.0000	2.00000	20
Total	4.2000	1.89737	40

Same group means that we've seen before.

Parameter Estimates

Dependent Variable: rating of depression -- bigger scores are poorer

Parameter	B	Std. Error	t	Sig.
Intercept	3.856	.135	28.536	.000
[grp=1.00]	.411	.191	2.153	.038
[grp=2.00]	0[a]	.	.	.
cov_c	-.107	.008	-13.871	.000
[grp=1.00] * cov_c	.205	.011	18.565	.000
[grp=2.00] * cov_c	0[a]	.	.	.

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

These values are the same as from the regression analysis.

Remember that SPSS dummy codes the grouping variable with the highest-valued group as the comparison group.

SPSS computes the interaction codes as the product of the dummy code of the grouping variable and of the centered continuous variable (you must remember to use the centered version of the variable when you submit the analysis).

Tests of Between-Subjects Effects

Dependent Variable: rating of depression -- bigger scores are poorer

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	127.332[a]	3	42.444	116.930	.000
Intercept	655.817	1	655.817	1806.717	.000
grp	1.682	1	1.682	4.634	.038
cov_c	.212	1	.212	.583	.450
grp * cov_c	125.114	1	125.114	344.677	.000
Error	13.068	36	.363		
Total	846.000	40			
Corrected Total	140.400	39			

a. R Squared = .907 (Adjusted R Squared = .899)

Using GLM in this way provides an F-test for each specific "effect" in the model.

The interaction $F = t^2$ from the interaction regression weight.

The group $F = t^2$ from the group reg wt.

The cov_c $F \neq t^2$ from the cov reg weight!!!

Why? GLM uses dummy coding (0,1) to compute the regression weight for grp, but it uses effect coding (-.5, .5) to compute the SS for grp.

So, when grp is dummy coded (in the parameter estimates) , the regression weight for cov_c tells the slope of the DV-cov_c regression line for those in the comparison group (coded 0).

But, when grp is effect coded (in the ANOVA table) the cov_c effect is testing the slope of the DV-cov_c regression line for those with grp = 0 (which is no one because of effects coding).

If the groups have the same n, then this F tests the "main effect" slope of the DV-cov_c regression line (i.e., "on average" for those coded -1 and 1).

treatment condition

Dependent Variable: rating of depression -- bigger scores are poorer

treatment condition	Mean	Std. Error
support group	4.406[a]	.135
cog-beh group	3.994[a]	.135

a. Covariates appearing in the model are evaluated at the following values: COV_C = .0000, INT = .6750.

Above are the group means after correction for the covariate and the interaction. These are the same values as on the respective regression lines for each group at COV = 0.

The difference between these corrected means is the same as the regression weight for the dummy-coded group term in both the Parameter Estimates from GLM and the regression weights from the regression analysis up above.