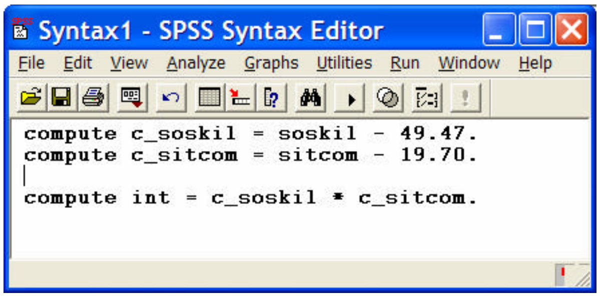
Regression Including the Interaction Between Quantitative Variables

The purpose of the study was to examine the inter-relationships among social skills, the complexity of the social situation, and performance in a social situation. Each participant considered their most recent interaction in a group of 10 or larger that included at least 50% strangers, and rated their social performance (**perf**) and the complexity of the situation (**sitcom**), Then, each participant completed a social skills inventory that provided a single index of this construct (**soskil**). The researcher wanted to determine the contribution of the "person" and "situation" variables to social performance, as well as to consider their interaction.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
SITCOM	60	.00	37.00	19.7000	8.5300
SOSKIL	60	27.00	65.00	49.4700	8.2600
Valid N (listwise	60				



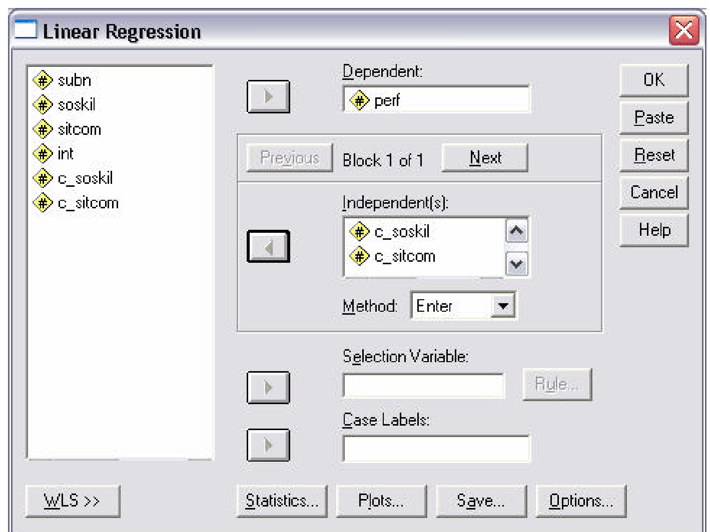
As before, we will want to center our quantitative variables by subtracting the mean from each person's scores.

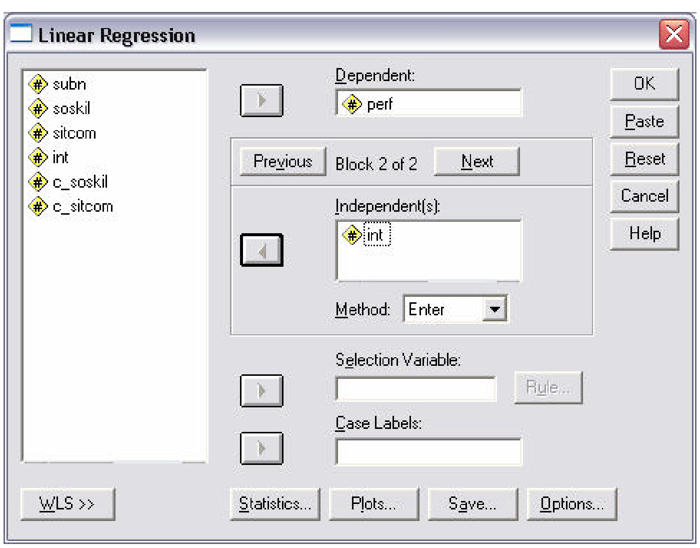
We also need to compute an interaction term as the product of the two centered variables.

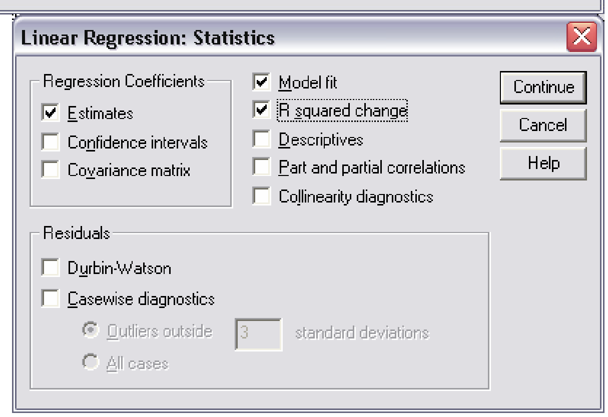
Some prefer using a "full model" approach, others a "hierarchical model" approach – remember they produce the same results.

Using the hierarchical approach, the centered quantitative variables (main effects) are entered on the first step and the interaction is added on the second step

Be sure to check the "R-squared change" on the Statistics window







SPSS Output:

Model 1 is the main effects model and Model 2 is the full model.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.374[a]	.140	.110	11.3406	.374	4.636	2	57	.014
2	.526[b]	.277	.238	10.4899	.137	10.620	3	56	.002

a. Predictors: (Constant), C_SITCOM, C_SOSKIL

b. Predictors: (Constant), C_SITCOM, C_SOSKIL, INT

The R Square Change for Model 2 is the contribution of the interaction. The F Change for Model 2 is the test of the interaction HO:.

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1192.355	2	596.178	4.636	.014[a]
	Residual	7330.742	57	128.610		
	Total	8523.097	59			
2	Regression	2360.961	3	786.986	7.152	.000[b]
	Residual	6162.137	56	110.038		
	Total	8523.098	59			

a. Predictors: (Constant), C_SITCOM, C_SOSKIL

b. Predictors: (Constant), C_SITCOM, C_SOSKIL, INT

c. Dependent Variable: PERF

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	29.102	1.464		19.878	.000
	C_SOSKIL	.524	.180	.360	2.914	.005
	C_SITCOM	.210	.174	.149	1.204	.234
2	(Constant)	29.651	1.365		21.728	.000
	C_SOSKIL	.378	.172	.260	2.196	.032
	C_SITCOM	.147	.162	.104	.907	.368
	INT	7.120E-02	.022	.385	3.259	.002

a. Dependent Variable: PERF

Inspection of the full model tells us that we have an interaction – the R^2 change = .137, which is nearly ½ of the variance accounted for by the model, and the R^2 change test tells us that increase is larger than would be expected by chance (the same information is provided by the test of the interaction **b** weight in the full model – t^2 - F-change).

Note: Raw regression weights for interactions are often numerically small. Why? Remember the interaction term is computed as the product of the centered main effect term. It is likely to have a very large standard deviation, and so, a much smaller regression weight than the main effects. Always use ß weights to consider the "relative contribution" of the model terms.

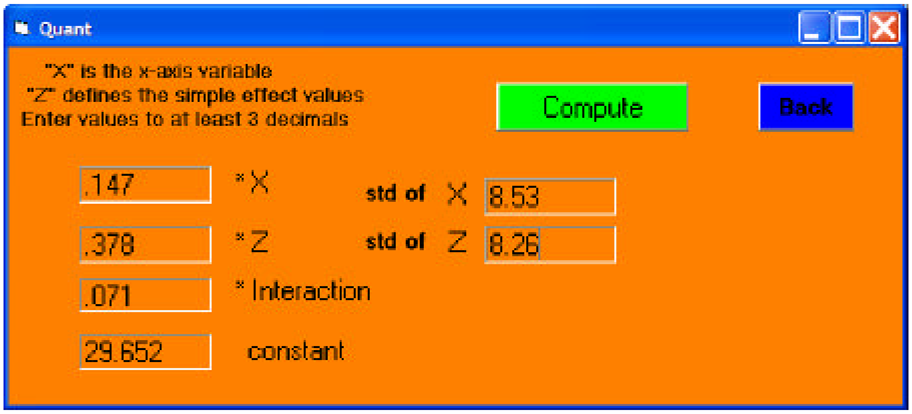
We can also see from the full model that there is a "main effect" for social skill, but not for the complexity of the social situation. We should be careful **not** to say that situational complexity does not contribute to this model!! Even though there is no main effect for this variable, it is involved in the interaction. Also, remember that because there is a significant interaction, we need to check carefully to determine if the main effects (significant or not) are descriptive or misleading.

Now that we have determined that there is an interaction between these variables, we need to identify the pattern of that interaction, as a basis for interpreting it. To do this we will plot the model, interpret it and "tell the story' of how these variables relate to the criterion.

One Small Difference when Plotting the Interaction of 2 Quantitative Variables

When we plotted the model for the interaction between a quantitiative variable and a categorical variable, we plotted the separate Y-X regression line for each of the different Z groups. Now, however, we have no groups – but we need some lines! The usual convention is to plot the Y-X regression line for the mean of Z, for +1 std above the mean of Z and -1 std below the mean of Z. This roughly corresponds to plotting the his roughly corresponds to plotting the Y-X relationship for small, medium and large values of Z. You can also plot the Y-X lines for any specific value of Z that you like.

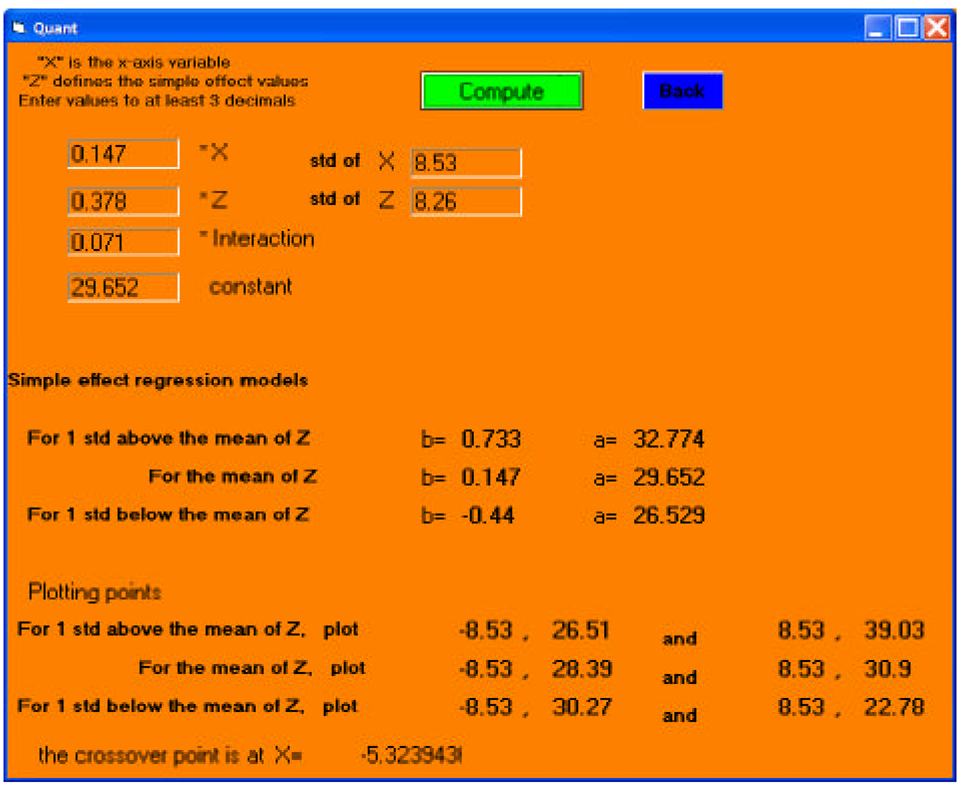
We will use the “2 quant predictors w/ +/- Std Z” portion of the IntPlot computator to obtain the plotting information.



I have chosen to examine the interaction by looking at the relationship between social performance and situational complexity, for different amounts of social skill, so

social complexity is “X”
social skill is “Z”

enter the regression weights from the full model and the standard deviation for each variable

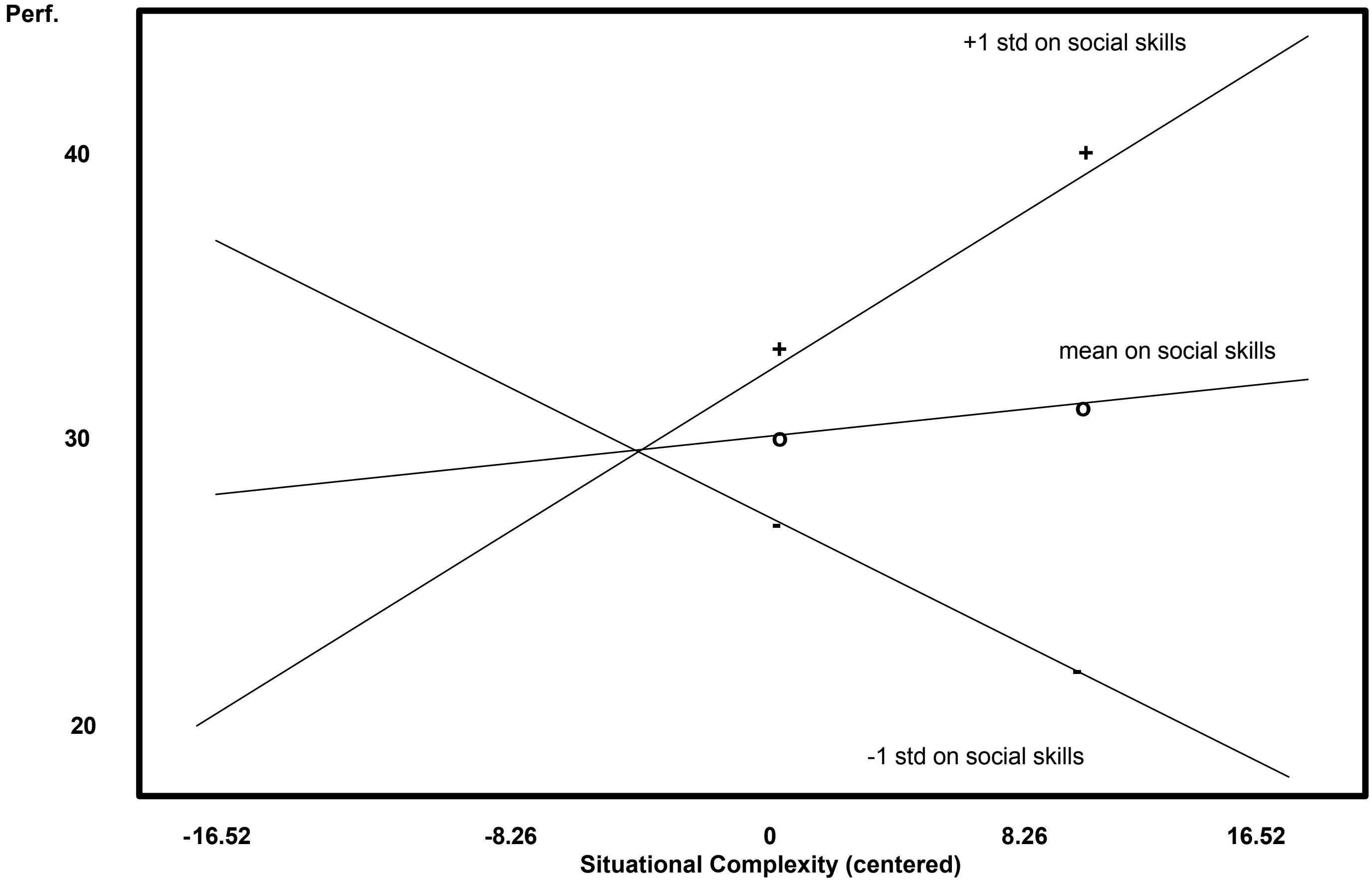


The output incudes the specific Y-X regression line for each value of Z

It aso gives ploting points for each of these lines

Finally, it gives the “cross over point” for where the three Y-X lines converge, which can make plotting the separate lines much easier

Plotting these simple regression lines, we get the following figure (ploting values are marked as -, o, and +, respectively).



Interpreting Regression Weights

Like with factorial designs, it is a good idea to start by interpreting the interaction (because whether or not it is significant is important do decide how to interpret the main effects)

the interaction b weight tells the direction and extent of the change in the slope of the Y-X regression line for a 1-unit increase in Z (or the direction and extent of the change in slope of the Y-Z regression line for a 1-unit change in X)

- a + weight means that as z gets larger, the Y-X slope gets more positive
- a 0 weight (no interaction) means that the Y-X slope is the same for all values of Z
- a – weight means that as Z gets larger, the Y-X slope gets less positive

the b weight of X tells the slope of the Y-X line at the mean of Z

the constant (a) tells the y-intercept for the Y-X line at the mean of Z (which was centered)

the b weight of Z tells how the direction and extent of the change in the height of the Y-X regression line for a 1-unit increase in Z

The crossover of the separate regression lines precludes having descriptive main effects!

- Since the regression lines vary from + to -, the Y-X regression line for the mean of Z isn't descriptive for all values of Z.
- Since the regressions line cross, differences between the height of the different lines at the 0-point of X isn't descriptive for al values of X

Using SPSS to get the Separate Regression Lines

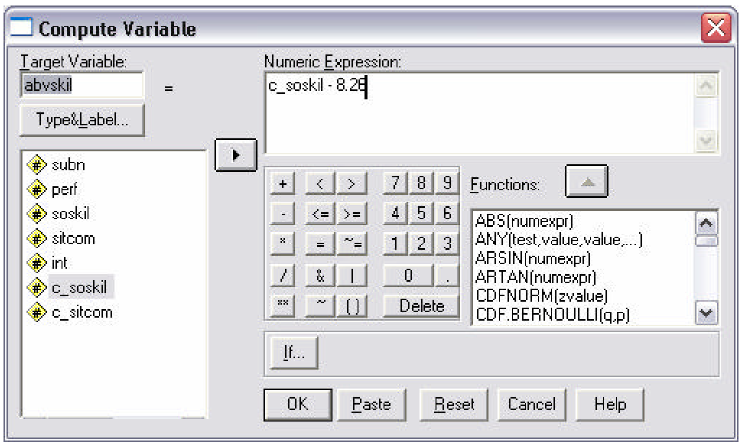
The draw-back to the approach we just took is that we do not have significance tests of whether the simple regression slopes are different from zero. When we want that formal test, we must apply the following procedure.

Obtaining the simple regression line for complexity 1 std ABOVE the mean of social skills

-- we need to "recenter" the scores around a point one standard deviation above the mean
-- this means, in effect, that we need to "lower" all the scores by one standard deviation
-- then we need to compute an interaction term specific to these "recentered" values
-- these new variables are used to obtain the simple regression line

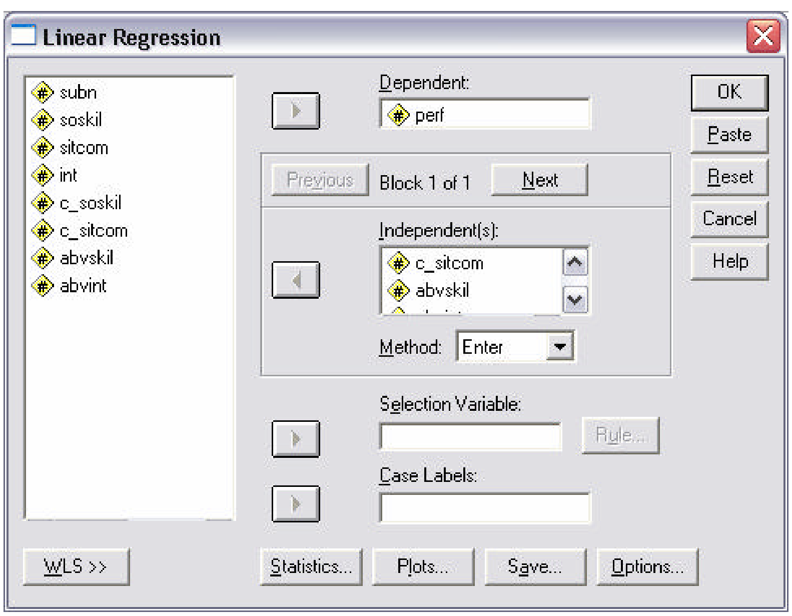
SPSS: Transform → Compute

Compute the variable abvskil by subtracting the standard deviation (8.26) from c_soskil.
Then compute the recentered interaction term abvint by multiplying abvskil by c_sitcom.



SPSS: Analyze → Regression → Linear

Enter perf as the dependent variable and c_sitcom, abvskil and abvint as the Independents.



SPSS Output:

Notice that the R^2 is the same as the original model ! - the recentering "re-partitions" the variance in the full model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.526[a]	.277	.238	10.4899

a. Predictors: (Constant), ABVINT, ABVSKIL, C_SITCOM

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2360.960	3	786.986	7.152	.000[a]
	Residual	6162.137	56	110.038		
	Total	8523.090	59			

a. Predictors: (Constant), ABVINT, ABVSKIL, C_SITCOM

b. Dependent Variable: PERF

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	32.775	1.930		16.896	.000
	C_SITCOM	.735	.228	.522	3.225	.002
	ABVSKIL	.378	.172	.260	2.196	.032
	ABVINT	7.120E-02	.022	.513	3.259	.002

a. Dependent Variable: PERF

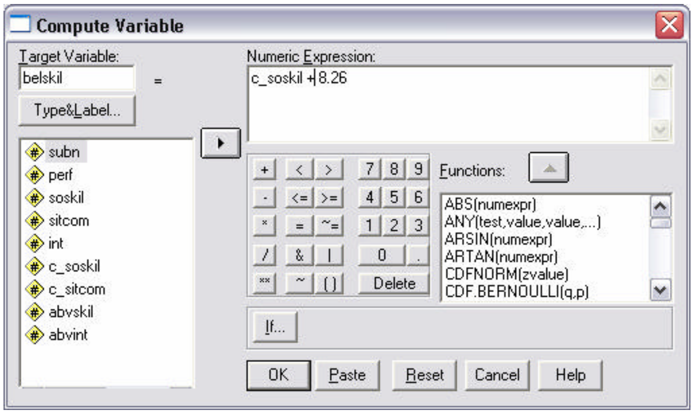
The b-weight and its t-test tells us that there is a significant, positive relationship between performance and social complexity for those 1 std above the mean on social skills (same value as was obtained by hand-calculation above). One could interpret this as meaning that, for those with particularly good social skills, increased social complexity "brings out their best".

Obtaining the simple regression line for complexity 1 std BELOW the mean of social skills

-- we need to "recenter" the scores around a point one standard deviation below the mean
-- this means, in effect, that we need to "raise" all the scores by one standard deviation
-- then we need to compute an interaction term specific to these "recentered" values
-- these new variables are used to obtain the simple regression line

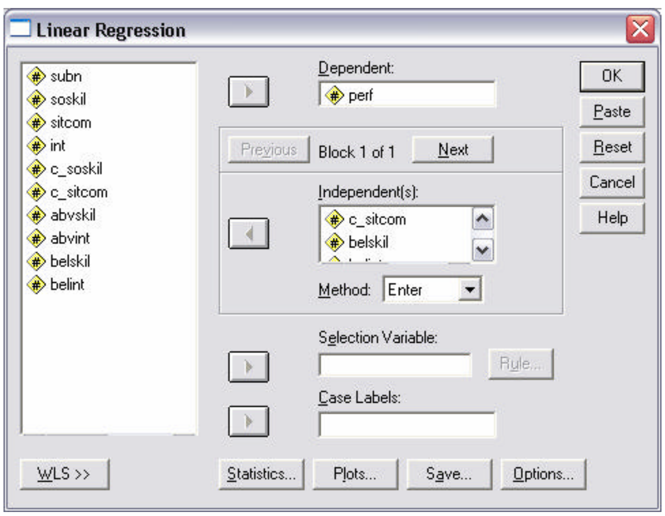
SPSS: Transform → Compute

Compute the variable belskil by adding the standard deviation (8.26) to c_soskil. Then compute the recentered interaction term belint by multiplying belskil by c_sitcom.



SPSS: Analyze → Regression → Linear

Enter perf as the dependent variable and c_sitcom, belskil and belint as the Independents.



SPSS Output:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.526[a]	.277	.238	10.4899

a. Predictors: (Constant), BELINT, BELSKIL, C_SITCOM

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2360.960	3	786.987	7.152	.000[a]
	Residual	6162.137	56	110.038		
	Total	8523.097	59			

a. Predictors: (Constant), BELINT, BELSKIL, C_SITCOM

b. Dependent Variable: PERF

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	26.529	2.002		13.250	.000
	C_SITCOM	-.441	.257	-.313	-1.719	.091
	BELSKIL	.378	.172	.260	2.196	.032
	BELINT	7.120E-02	.022	.592	3.259	.002

a. Dependent Variable: PERF

The b-weight tells us that there is a negative relationship between performance and social complexity for those 1 std below the mean on social skills (same value as was obtained by hand-calculation above). One could interpret this as meaning that, for those with particularly poor social skills, do not “handle” increased social complexity well, and their performance suffers.

You will probably notice that the significance test of the simple regression weight we have just interpreted does not reject H0:. What to do? Remember that the choice of +/- 1 standard deviation is arbitrary, so many folks who do this sort of analysis suggest re-computing the simple slopes at +/- 1.5 standard deviations. It is likely that this more extreme set of simple slopes will be “more different” from the slope for the mean of social skills.