QxQ Models: Using Regression & GLM for Linear Models Including Interactions

The purpose of the study was to examine the contribution of the "person" and "situation" variables to social performance, as well as to consider their interaction. Specifically, the study was designed to test the inter-relationships among social skills, the complexity of the social situation, and performance in a social situation. Sixty participants who had were selected for having a wide range of dyadic and group social skills, based on a previously completed survey, were each invited to one of 15 psychology club get-to-know-you parties. Each party had a manipulated mix of formally and informally dressed male and female, administrators, faculty, undergraduates and graduates. The mix was designed to manipulate the "social complexity" of the situation according to a scoring system previously devised and validated. Also at each meeting there were several teams of trained research confederates who engaged each participant in a conversation that was carefully scripted and staged. Video of these conversations were coded to provide a social performance score for each participant.

Regression: Basic Model

```
*centering each of the quantitative predictors.
*computing the interaction.
compute c_soskil = soskil – 49.4700.
compute c_sitcom = sitcom – 19.7000.
compute int = c_soskil * c_sitcom.

REGRESSION
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /DEPENDENT perf
  /METHOD=ENTER c_soskil  c_sitcom
  /METHOD=ENTER int.
```

This shows the hierarchical approach to obtaining and testing this model – obtaining first the main effects model, and then testing if adding the interaction significantly increases the fit of the model.

← **"CHANGE"** gets the $R^2\Delta$ F-test

← requests the main effects model

← adds the interaction to obtain the full model

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

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.374[a]	.140	.110	11.3406	.140	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

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

Both the main effects model and the full model "work".

$R^2\Delta$ for Model 2 = .137. More than ½ of the variance accounted for by the full model is due to the interaction

We have 2 equivalent tests of the interaction:

The $R^2\Delta$ F-test for Model 2 tells us this $R^2\Delta$ is significant.

The t-test of the interaction **b** weight in the full model tells us that the contribution of the interaction is significant.

t^2 (3.259^2) = $R^2\Delta$ F (10.62)

Coefficients [a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
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

Interpreting the Regression Weights from the Main Effects Model

When the full model fits the data significantly better than does the main effects model, it is usually a good idea not to spend too much time interpreting the main effects model; the main effects are likely to be misleading.

Main Effects *"versus"* Full Models – *Two Important Differences*

Most multiple regression models are "main effects models" – involving no interactions. So, it is important to remember how the interpretation of a regression weight is different when the model includes an interaction.

As you remember the effects in main effects models are "main effects" (the relationship between that variable and the criterion, after controlling for the other effects in the model), while the effects in a full models are "simple effects" (the relationship between that variable and the criterion controlling for the other variables at "0" – the mean after mean-centering). But also remember, that the effects in a main effects model are controlled (only) for each other, while the effects in a full model are also controlled for each other and also controlled for the interaction. Including this "other predictor" could produce any of the "multivariate" effects we have come to expect, especially when we consider that the collinearity between predictors and their interaction are usually higher than among individual predictors. So, suppressor & collinearity effects should be considered and interpreted carefully.

Interpreting the Regression Weights from the Full Model (Including the Interaction)

Simple effect for social skills when situational complexity = 0 (its mean after centering)

C_SOSKIL -- The regression weight for social skills tells that for average situational complexity, there is a positive relationship between social skills and performance. This positive slope is statistically significant.

Simple effect for the situational complexity when social skills = 0 (its mean after centering)

C_SITCOM -- The regression weight for situational complexity tells us that for those with average levels of social skills, there is not a significant relationship between situational complexity and performance.

Interaction – simple effect of each predictor is different for different levels of the other variable

INT -- The interaction regression weight tells us that, both

- For each 1-unit increase in social skills, the slope of the relationship between performance and situational complexity increases by .007 (this increase in slope is statistically significant)
- For each 1-unit increase in situational complexity, the slope of the relationship between performance and social skills increases by .007 (this increase in slope is statistically significant)
- **Note:** Raw regression weights for interactions are often numerically small. Why? The interaction term is computed as the product of the centered main effect terms. So it will have a relatively large standard deviation and consequently a relatively small regression weight compared to the main effects. Always use β weights to consider the "relative contribution" of the main effects and interaction.

Plotting the QxQ Model

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 qualitative predictor groups. Now, however, we have no groups – but we still need some lines! The xls plotting Computator follows the usual convention, which is to plot the Y-X regression line for the mean of the 2^{nd} quant predictor, for +1 std above the mean, and -1 std below the mean.

You have to decide which quantitative predictor to put on the x-axis. Much like the decision for deciding how to compose tables or figures for factorial designs, you should put your “primary effect” on the x-axis, so that you get to see how the relationship between this primary effect and the DV differs for different values of the other variable.

Remember – all the analyses shown in the following pages produce the same model!!! We may recode this or re-center that to change the specific information available from a regression weight and a significance test. We may even change which variable is shown on the x-axis. No matter what, they are all the same model!

Plotting the QxQ Model

Usually we don’t plot main effects models. If there are no interactions, then a careful interpretation of each of the main effect regression weights provides a complete description of the model. However, when we have a contributing interaction term in the model, the weights are a collection of simple effects and simple effect differences, that are not a complete description of the model. So, we plot the interaction model to get a more complete picture and description (usually in combination with a collection of re-centered analyses to get specific targeted significance tests – more below)

height x1-=mean x2=mean	constant	29.102
y-x1 slope for x2=mean	b(x1)	0.147
y-x1 height dif for dif x2	b(x2)	0.378
y-x1 slope dif for dif x2	b(x1x2)	0.071
Situational Complexity	x1(mean)	19.7
	x1(std)	8.53
Social Skills	x2(mean)	49.47
	x2(std)	8.26

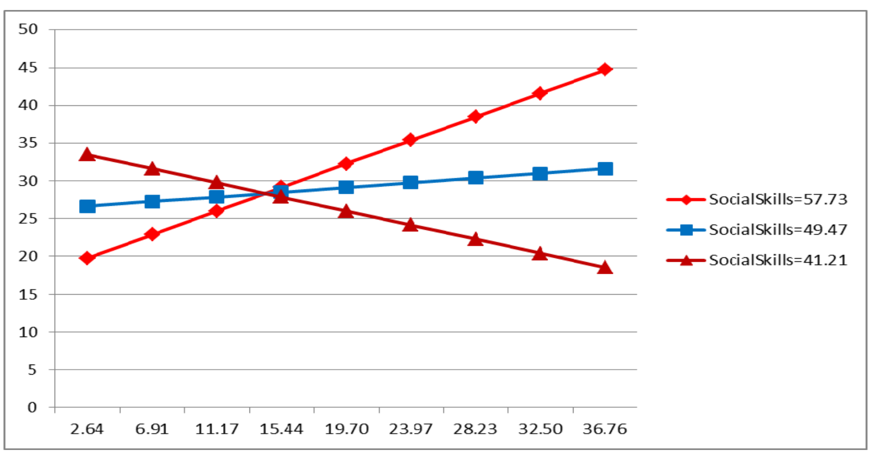
← Fill in the regression weights – be sure to get “x1” & “x2” right

← use “0” for the interaction weight if you have a main effects model

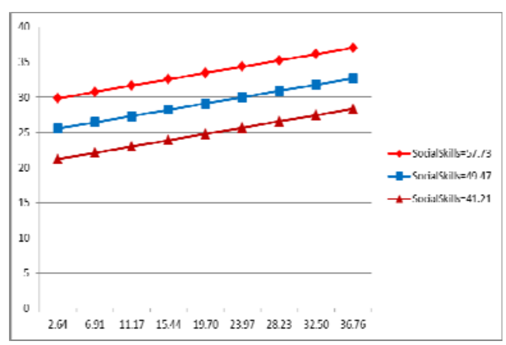
← Fill in the mean & std for each quant variable – remember “x1” is the variable you intend to be on the x-axis of the model plot.

	(slope * X) +		height
SocialSkills=57.73	0.73346	*X +	32.22428
SocialSkills=49.47	0.147	*X +	29.102
SocialSkills=41.21	-0.4395	*X +	25.97972

The program will generate the simple regression slope for 3 values of the other variable: -1 std, mean & +1` std



Here’s the main effects model –for comparison.



The diffferences between the main effect and full models are drammatic.

- The negative performance-complexity slope for those with low social skills in the full model is very different from the positive slope in the main effect model.
- The inversion of the performance-social skills relationship at low values of complexity is also very different from what we see in the mail effects model.

Using SPSS to get & test Specific Regression Lines

At this point we know that the Performance-Situational Complexity regression line is not different from 0 for those at the mean level of Social Skills (sitcom regression weight). We also know that the slope of the Performance-Situational Complexity regression line varies with the level of Social Skills (i.e, the significant interaction). However, we do not have significance tests to tell us if the slopes of the Performance-Situational Complexity regression lines are significantly different from 0 for any specific social skills values. When we want that formal test, we can apply the following procedure.

Obtaining the performance-complexity simple regression line for +1 std social skills

- we need to "re-center" 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 "re-centered" values

```
*re-center social skills at +1 std.
*use the original mean-centered complexity variable.
*computing the new interaction term.
compute abvskil = c_soskil – 8.26.
compute c_sitcom = sitcom – 19.7000.
compute abvint = abvskil * c_sitcom.

REGRESSION
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /DEPENDENT perf
  /METHOD=ENTER abvskil  c_sitcom  abvint
```

The R^2 & F-test results will be the same as the original model. We did this analysis just to get the regression weight for Situational Complexity with Social Skills re-centered.

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 situational complexity variable tells us that there is a significant, positive relationship between performance and situational complexity for those 1 std above the mean on social skills (same value as was obtained from the computator above).

One could interpret this as meaning that, for those with particularly good social skills, increased social complexity "brings out their best".

Obtaining the performance-complexity simple regression line for -1 std 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

```
*re-center social skills at -1 std.
*use the original mean-centered complexity variable.
*computing the new interaction term.
compute belskil = c_soskil + 8.26.
compute c_sitcom = sitcom – 19.7000.
compute belint = belskil * c_sitcom.

REGRESSION
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /DEPENDENT perf
  /METHOD=ENTER belskil  c_sitcom  belint
```

SPSS Output:

The R^2 & F-test results will be the same as the original model. We did this analysis just to get the regression weight for Situational Complexity with Social Skills re-centered.

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 and its t-test situational complexity variable tells us that there is a nonsignificant, negative relationship between performance and situational complexity for those 1 std below the mean on social skills (same value as was obtained from the computator 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:. Is this a problem?

Remember that the choice of +/- 1 standard deviation is arbitrary. One common approach is to try re-dentering at multiple different social skill values, to find the range within which the performance-complexity relationship is positve, non-significant and negative.