

## 2xQ Models: Using Regression for Linear Models Including Interactions

The purpose of this study was to examine the relationship of practice and exam performance. Two aspects of practice were selected for study, the difficulty of the practice and the number of practices completed. Practice difficulty was a 2-condition variable - practice problems were either about the same difficulty as the exam problems (=1) or they were easier than the exam problems (=2). Different sections of the course were randomly assigned to receive the two difficulty levels. Students were permitted to complete as many practice problems as they liked, receiving very complete feedback after each problem. The dependent variable was performance on an examination.

Here are the group means for number of practices and exam performance.

There is a practice difficulty group difference for number of practices – those in the easier condition completed fewer practices.

There is also a practice difficulty group difference for test performance.

Notice the confounding – the group with the higher average number of practices is also the group with the higher average performance – confounding our comparison of the same and easier conditions.

Report

practgrp		numpract	testperf
same	Mean	6.6250	67.5000
	N	16	16
	Std. Deviation	2.60448	12.90994
easier	Mean	5.2500	57.5000
	N	16	16
	Std. Deviation	2.62043	8.56349
Total	Mean	5.9375	62.5000
	N	32	32
	Std. Deviation	2.66322	11.91367

### Data Preparation

Here's the SPSS syntax code to dummy code the binary grouping variable, to center the quantitative variable (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 simplifies interpretation of the regression weights
- Interactions are “non-additive combinations” -- meaning products of the related main effects

```
* pract_dc1 compares same=1=>1 with easier = 2=>0.  
if (practgrp = 1) pract_dc = 1.  
if (practgrp = 2) pract_dc = 0.
```

```
compute numpract_cen = numpract - 5.938.
```

```
compute grp_pract_int = pract_dc * numpract_cen.
```

```
exe.
```

IF statements to dummy-code the group variable: same is coded “1” as the target group and easier is coded “0” as the comparison group

Centering the covariate requires subtracting the mean from each person's number of practices score

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

## Main effects model or ANCOVA assuming regression slope homogeneity (no interaction)

regression dep testperf  
/ enter numpract\_cen pract\_dc

← identifies the criterion variable  
← identifies the predictors

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.493 <sup>a</sup>	.243	.191	10.71622

a. Predictors: (Constant), numpract\_cen, pract\_dc

ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1069.719	2	534.860	4.658	.018 <sup>a</sup>
	Residual	3330.281	29	114.837		
	Total	4400.000	31			

a. Predictors: (Constant), numpract\_cen, pract\_dc  
b. Dependent Variable: testperf

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	58.290	2.728		21.366	.000
	pract_dc	8.422	3.926	.359	2.145	.040
	numpract_cen	1.148	.749	.257	1.533	.136

a. Dependent Variable: testperf

The model “works” significantly better than chance.

This model accounts for about 24% of the variance in the performance scores.

## Interpreting the regression weights

constant

- The expected value of testperf when the value of all predictors = 0
- The expected value of testperf for those in the easier condition and who have 5.938 practices
- Those in easier condition who had 5.938 practices scored 58.29% on the exam

numpract\_cen

- The direction and extent of the expected change in testperf for a 1-unit increase in this predictor, holding the value of the other predictor constant at 0
- The expected change in testperf for each additional practice, for those in the easier condition
- For each additional practice, those in the easier condition are expected to increase 1.148 % on testperf – this effect is not significant
- Note: because there is no interaction term in the model, and, so, the slopes of the two group's lines must be the same, this is also the slope of the relationship between practice and performance for the same condition

pract\_dc

- The direction and extent of the group difference, holding the other predictor constant at 0.
- The group difference controlling the number of practices at 0 (the mean after centering)
- Those in the same group outperformed those in the easier condition by 8.422, when holding for the number of practices at 5.938 – this effect is significant.
- So, the corrected mean for the easier condition when practice is controlled at is 58.29% (constant) and corrected mean difference between the groups is 8.422% (group regression weight), so the corrected mean for the same condition is 66.712%
- Notice that this corrected group difference is smaller than the 10-point uncorrected group difference between the groups (67.5 vs. 57.5).

## Obtaining & Interpreting the Plot of the Model

<i>height z=0</i>	constant	58.29	
<i>slope z=0</i>	b(x)	1.148	Practice
<i>height dif z=1</i>	b(z)	8.422	Easier
<i>slope dif z=1</i>	b(xz)	0	Same
	x(mean)	5.938	
	x(std)	2.663	

Using the "2xQ Linear" tab of the Excel file...

Label the groups -- be sure you label correctly!

Fill in the values from the analyses.

Put "0" in for the interaction regression weight.

	( slope * X ) +	height
Easier	1.148 * X +	58.29
Same	1.148 * X +	66.712

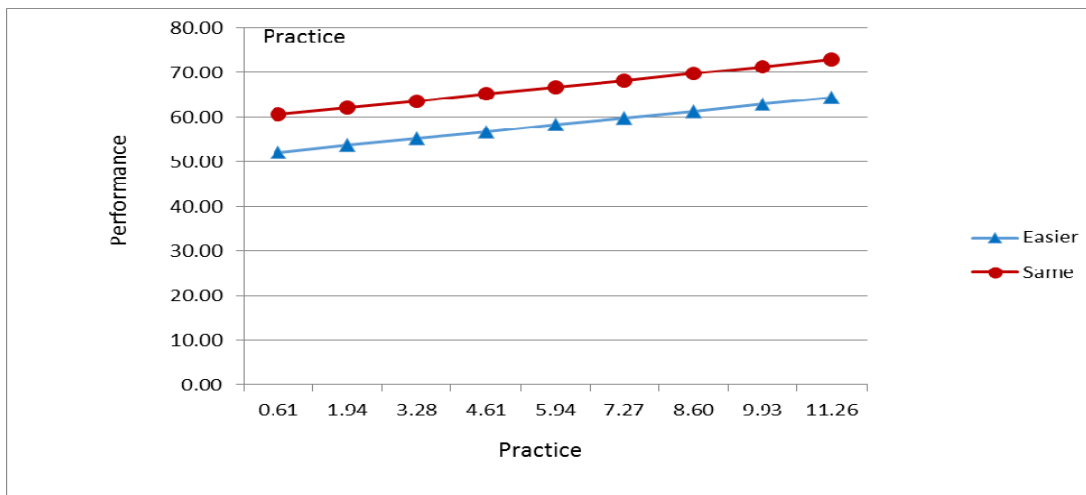
The weights for the simple regression line for each group are automatically calculated.

Notice the slope of the simple regression line for each group is the same – with no interaction term in the model we are forcing a homogeneity of regression slope assumption onto the model.

The only difference between the two simple regression lines is their height, which is given by the regression weight for the grouping variable.

x std range		-2	-1.5	-1	-0.5	0	0.5	1	1.5	2
	x-centered	-5.33	-3.99	-2.66	-1.33	0.00	1.33	2.66	3.99	5.33
z wt	x-row	0.61	1.94	3.28	4.61	5.94	7.27	8.60	9.93	11.26
0	y'	52.18	53.70	55.23	56.76	58.29	59.82	61.35	62.88	64.40
1	y'	60.60	62.13	63.65	65.18	66.71	68.24	69.77	71.30	72.83

The plotting points for the graph are automatically computed. Both the raw and the centered x-axis values are shown.



Because there is no interaction term, the main effects are "safe". So...

The significant group difference when controlling at 0 (the mean number of practices after mean centering) generalizes into a group difference at all amounts of practice

The null relationship between practice and performance for the group coded 0 (Same difficulty) generalizes to both groups.

## Full model - including the interaction

regression dep testperf

← identifies the criterion variable

/ enter numpract\_cen pract\_dc grp\_pract\_int

← identifies the predictors

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.674 <sup>a</sup>	.455	.396	9.25813

a. Predictors: (Constant), grp\_pract\_int, pract\_dc, numpract\_cen

ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2000.036	3	666.679	7.778	.001 <sup>a</sup>
	Residual	2399.964	28	85.713		
	Total	4400.000	31			

a. Predictors: (Constant), grp\_pract\_int, pract\_dc, numpract\_cen  
b. Dependent Variable: testperf

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	56.832	2.398		23.699	.000
	pract_dc	8.406	3.392	.358	2.478	.020
	numpract_cen	-.971	.912	-.217	-1.064	.296
	grp_pract_int	4.263	1.294	.660	3.295	.003

a. Dependent Variable: testperf

The model “works” significantly better than chance.

This model accounts for about 46% of the variance in the performance scores.

## Interpreting the regression weights

constant

- The expected value of testperf when the value of all predictors (practice difficulty, number of practices and interaction) = 0
- The expected value of testperf for those in the easier condition and who have 5.938 practices
- Those in easier condition who had 5.938 practices scored about 58.83% on the exam

numpract\_cen

- The direction and extent of the expected change in testperf for a 1-unit increase in this predictor, holding the value of the other predictors (practice & interaction) constant at 0
- The expected change in testperf as the number of practice changes for those in easier condition
- For each additional practice, those in the easier condition are expected to decrease .971 % on testperf – this effect is not significant
- Note: because there is an interaction term in the model, the slopes of the two group’s lines may be different – check the interaction to evaluate this.

pract\_dc

- The direction and extent of the testperf group difference, holding the other predictors (number of practices and the interaction constant at 0.
- Group difference controlling the number of practices and interaction at 0 (mean after centering)
- Those in the same group outperformed those in the easier condition by 8.406, when holding for the number of practices at 5.938 – this effect is significant.
- So, the corrected mean for the easier condition when practice is controlled at is 56.832% (constant) and corrected mean difference between the groups is 8.406% (group regression weight), so the corrected mean for the same condition is 65.238%
- Notice that this corrected group difference the mean # practices is smaller than the 10-point uncorrected group difference between the groups (67.5 vs. 57.5) – the effect is smaller after correcting for the confounding # practices.

grp\_pract\_int

- The direction and extent of the difference in the testperf-numpract slope for the two groups.
- The direction and extent of change in the practice difficulty group difference for each 1-unit increase in number of practices
- How the practice group difficulty effect changes as the number of practices changes
- For each additional practice, the difference between the similar difficulty practice group and the easier practice group increases by 4.264% – this effect is significant.
- So, for those in the easier practice group performance decreases by .971% for each practice, whereas for those in similar difficulty group, performance increases by 3.292% (-.971 + 4.263).

## Obtaining & Interpreting the Plot of the Model

height z=0	constant	56.832		
slope z=0	b(x)	-0.971	Practice	z wt
height dif z=1	b(z)	8.406	Easier	0
slope dif z=1	b(xz)	4.263	Same	1
	x(mean)	5.938		
	x(std)	2.663		

Using the “2xQ Linear” tab of the Excel file...

Label the groups -- be sure you label correctly!

Fill in the values from the analyses

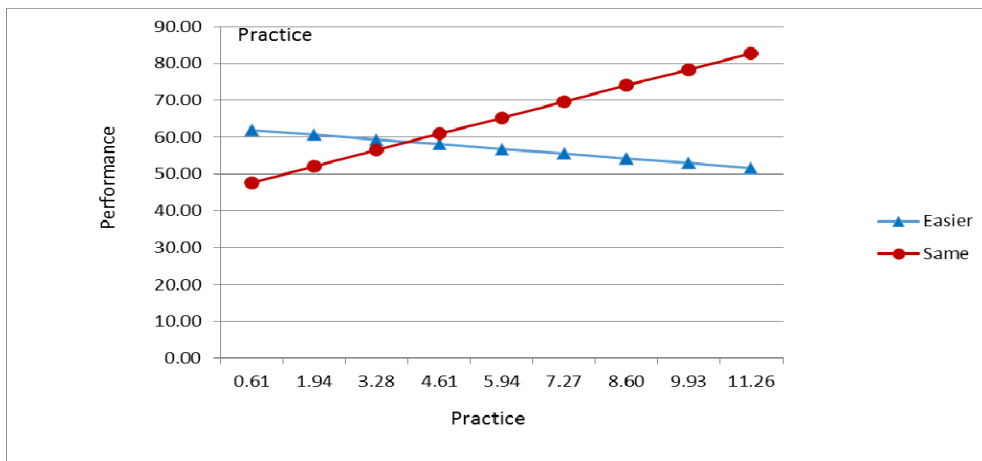
The weights for the simple regression line for each group are automatically calculated.

Notice the slope of the simple regression lines for the two groups are different – by the about of the interaction regression weight!

	(slope * X) +	height
Easier	-0.971 * X +	56.832
Same	3.292 * X +	65.238

x std range		-2	-1.5	-1	-0.5	0	0.5	1	1.5	2
	x-centered	-5.33	-3.99	-2.66	-1.33	0.00	1.33	2.66	3.99	5.33
z wt	x-row	0.61	1.94	3.28	4.61	5.94	7.27	8.60	9.93	11.26
0	y'	62.00	60.71	59.42	58.12	56.83	55.54	54.25	52.95	51.66
1	y'	47.70	52.09	56.47	60.85	65.24	69.62	74.00	78.39	82.77

The plotting points for the graph are automatically computed. Both the raw and the centered x-axis values are shown.



We can see the pattern of the data

For those in the Easier condition (coded 0) practice is not related to performance (-.971,  $p=.296$ ).

The significant positive interaction term tells us that the slope is significantly more positive (4.263 more,  $p = .003$ ) for the Same group (coded 1), than for the Easier group – but we do not have a significance test of the slope for the Same group.

So, it is likely that the main effect of practice is misleading.

At the average number of practices (0 after mean-centering) the Same Difficulty group scores significantly higher (8.406,  $p = .020$ ) than Easier Practice group. We do not have group comparison tests for any other values of practice. However, we can see that the group effect varies from one direction to the other, rendering the main effect for group misleading.