

## k-Group ANCOVA

The purpose of the study was to compare the Test Performance of students who had prepared for the test using practice problems that were: 1) easier than the test items (Easier) , 2) similar in difficulty to the test problems (Similar), and 3) harder than the test items (Harder). Students were randomly assigned to one of the practice groups and given a packet that had instructions for the problems, 10 practice problems (of the appropriate difficulty) and a set of 5 test problems. Students read the instructions, completed as many practice problems as they liked, and then completed the test problems. Practice group (practgrp), the number of practices completed (numpract) and test performance (testperf as a %) were recorded for each student.

Here are the results of ANOVAs comparing the groups on testperf and numpract.

### Descriptives

		N	Mean	Std. Deviation
testperf	easier	16	50.6250	12.89380
	same	16	67.5000	12.90994
	harder	16	89.6875	5.90727
	Total	48	69.2708	19.46053
numpract	easier	16	5.2500	2.62043
	same	16	6.6250	2.60448
	harder	16	8.0625	1.91377
	Total	48	6.6458	2.62143

There were significant test performance differences among the groups

There were also significant differences in the number of practices the students chose to complete.

This would seem to be an “augmenting confound” because the groups expected to perform better had the higher mean on the potential confound.

### ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
testperf	Between Groups	12282.292	2	6141.146	50.089	.000
	Within Groups	5517.188	45	122.604		
	Total	17799.479	47			
numpract	Between Groups	63.292	2	31.646	5.484	.007
	Within Groups	259.688	45	5.771		
	Total	322.979	47			

### Correlations

		testperf	numpract
testperf	Pearson Correlation	1	.416**
	Sig. (2-tailed)		.003
	N	48	48
numpract	Pearson Correlation	.416**	1
	Sig. (2-tailed)	.003	
	N	48	48

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Similarly, while there is a significant correlation between number of practice and test performance.

The positive correlation between number of practices and test performance supports the idea that this is an augmenting confound. The positive correlation suggests that whatever group had more practices will have a higher test performance.

Together, the group mean difference of number of practices and the correlation between number of practices and test performance suggest that it might be useful to perform an ANCOVA on these data.

## Data Preparation – Mean-centering the Covariate

It is a good idea to work with “mean-centered” quantitative covariate scores. Mean-centering simplifies the math involved in constructing and plotting the results of the analysis, as well as limiting collinearities among the models terms that can lead to mis-estimation and statistical conclusion errors.

Mean-centering is just what it sounds like... You compute a new variable for each person that is their covariate score minus the mean of covariate.

```
compute numpract_cen = numpract - 6.6458.  
exe.
```

## “Kinds” of ANCOVA models

Even for this, the simplest type of ANCOVA with a 2-group IV and a single covariate, there are different possible models.

**Main Effects ANCOVA** models include the IV and the Covariate. A main effects model makes the “homogeneity of regression slope” assumption. That is, the model is constructed assuming that the slope of the linear relationship between the covariate and the DV is the same for all IV groups. Put differently, this is an assumption that there is no interaction between the covariate and the IV as they related to the DV. This regression slope homogeneity assumptions makes the comparison of the IV groups simpler, in that, it assumes that the corrected mean DV difference between the groups is the same for all values of the covariate. In terms of this example, the assumption is that the test performance difference between the Easy, Similar and Harder difficulty practice groups is the same for every amount of practice.

**Full Model ANCOVA** models include the IV, the Covariate, and the IV-Covariate interaction. This model does not make the homogeneity of regression slope assumption, and allows there to be different corrected mean DV difference between the groups for different values of the covariate. Just like with factorial ANOVA, often the most important part of the model is the interaction! Also, sometimes, without careful attention to the pattern of the interaction, one or both main effects are misleading.

## Getting the Main Effects ANCOVA Model

Some of the useful output isn’t available using the SPSS GUI, so we will use SPSS syntax code for these analyses. The simplest code for an ANCOVA is shown below.

```
UNIANOVA testperf BY practgrp WITH numpract_cen    ← dv BY iv WITH covariate  
                                                    ← be sure to use the mean-centered cov  
  
/METHOD = SSTYPE(3)                                ← uses formulas that work well with ≈n  
  
/EMMEANS = TABLES(practgrp)                       ← gets dv means for each group  
              WITH(numpract_cen= mean)             ← corrected for the mean covariate value  
              COMPARE (practgrp)                  ← gets simple effects test for that cov value  
  
/PRINT = DESCRIPTIVE PARAMETER                     ← gets descriptive/uncorrected means and  
                                                    the regression model parameters (we will  
                                                    use to plot the model)  
  
/DESIGN = practgrp numpract_cen.                   ← specifies that the IV and the Covariate  
                                                    are both in the model (notice the period)
```

## Main Effects ANCOVA output

### Descriptive Statistics

Dependent Variable: testperf

practgrp	Mean	Std. Deviation	N
easier	50.6250	12.89380	16
same	67.5000	12.90994	16
harder	89.6875	5.90727	16
Total	69.2708	19.46053	48

### Tests of Between-Subjects Effects

Dependent Variable: testperf

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	12335.740 <sup>a</sup>	3	4111.913	33.114	.000
Intercept	230325.420	1	230325.420	1854.832	.000
practgrp	9254.383	2	4627.191	37.263	.000
numpract_cen	53.448	1	53.448	.430	.515
Error	5463.739	44	124.176		
Total	248125.000	48			
Corrected Total	17799.479	47			

a. R Squared = .693 (Adjusted R Squared = .672)

### Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	89.045	2.953	30.153	.000
[practgrp=1.00]	-37.787	4.394	-8.600	.000
[practgrp=2.00]	-21.535	4.063	-5.300	.000
[practgrp=3.00]	0 <sup>a</sup>	.	.	.
numpract_cen	.454	.692	.656	.515

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

These are the same (uncorrected) performance means we got from the ANOVA.

The F table shows that we have a significant practgrp effect after controlling for number of practices.

There is not a significant relationship between numpract and testperf, after taking group membership into account.

Notice that the SSerror is not much smaller in this ANCOVA model than in the original ANOVA model, telling us that the covariate didn't add much to the model, and probably won't change the group comparison much.

The parameter estimates are another "expression" of the information in the F table, but presented as t-tests of the multiple regression weights. We will primarily use this table to plot the model.

Because of the homogeneity of regression slope assumption the regression weight for the numpract\_cen tells the slope for all three groups.

For each t-test, df = 44, the dferror from the ANOVA model just above.

## Main Effects Model Corrected Means and their Comparison

### Estimated Marginal Means

#### practgrp

##### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
easier	51.258 <sup>a</sup>	2.948	45.316	57.200
same	67.509 <sup>a</sup>	2.786	61.895	73.124
harder	89.045 <sup>a</sup>	2.953	83.093	94.996

a. Covariates appearing in the model are evaluated at the following values: numpract\_cen = .0000.

##### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	-16.251 <sup>*</sup>	4.053	.000
	harder	-37.787 <sup>*</sup>	4.394	.000
same	easier	16.251 <sup>*</sup>	4.053	.000
	harder	-21.535 <sup>*</sup>	4.063	.000
harder	easier	37.787 <sup>*</sup>	4.394	.000
	same	21.535 <sup>*</sup>	4.063	.000

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

##### Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	9254.383	2	4627.191	37.263	.000
Error	5463.739	44	124.176		

The F tests the effect of practgrp. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

When the value of the covariate is held constant at its mean ("0" because of mean-centering) the estimated testperf is 89.045 for the Harder group, 67.509 for the Same Difficulty group and 51.258 for the Easier group.

The F-table tell us that these means are statistically significant.

The pairwise comparisons tell us that the three groups are significantly different from each other with Easier < Same < Harder

All, in all, the ANCOVA didn't tell us much more than did the ANOVA.

## Plotting the Main Effects ANCOVA Results

We will use an Excel plotting program for this. This uses the “kxQ Linear” tab, with the info for this analysis filled in. Be sure to:

- Change the name of the IV and include the IV group names
- Include the regression parameters from that table – put “0” for the interactions
- Include the mean and standard deviation of the covariate

You should also change the text boxes describing the Y- and X-axes of the plot.

Please note:

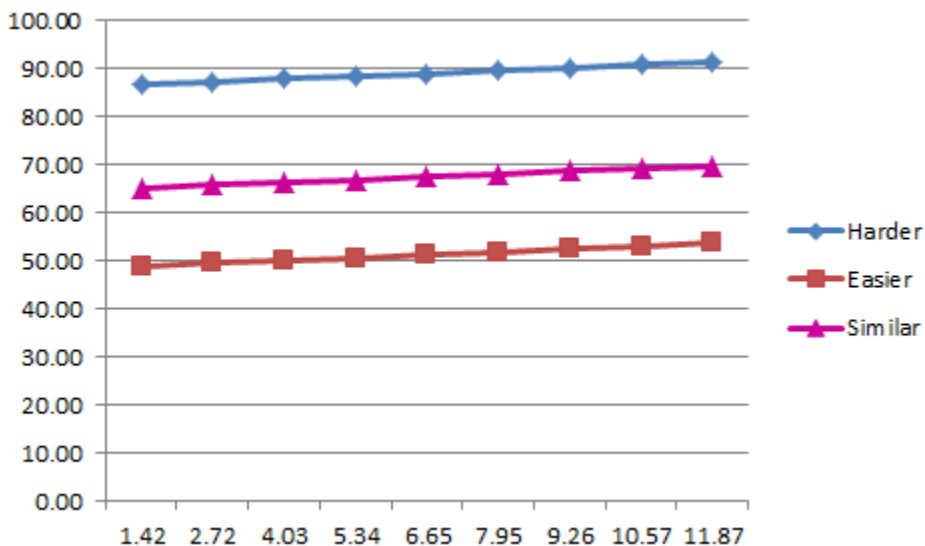
5	Constant	constant	89.045			
6	Covariate	b(x)	0.454	Practice	z1 wt	Z2 wt
7	IV/Groups - 1st	b(z1)	-37.787	Harder	0	0
8	Interaction -- 1st	b(xz1)	0	Easier	1	0
9	IV/Groups - 2nd	b(z2)	-21.535	Similar	0	1
10	Interaction -- 2nd	b(xz2)	0			
11						
12		x(mean)	6.6458			
13		x(std)	2.6143			

Use the IV group originally coded as “1” as the “z1 wt = 1 & z2 wt = 0” group (Easier difficulty in this example)

Use the IV group originally coded as “2” as the “z1 wt = 0 & z2 wt = 1” group (Similar difficulty in this example)

Use the IV group originally coded as “3” as the “z1 wt = 0 & z2 wt = 0” group (Harder difficulty in this example)

The program makes a pretty decent plot of the results...



The graph corresponds with the results from the F-table.

- The Harder Difficulty group had the best performance, and the Same Difficulty group did better than the Easier group
- The regression line relating Practice with Performance is pretty flat (notice that the regression lines for the groups are parallel → because of the regression slope homogeneity assumption)

## Getting the Full Model ANCOVA

There only a couple of differences when asking SPSS for the full model ANCOVA including the interaction term.

First, you will include the interaction term in the "DESIGN" subcommand. Represent this by listing the IV and Covariate, with "\*" between them → numpract\_cen\*practgrp (be sure to use the centered covariate)

Second, since the model allows for an interaction, and the slopes of the regression lines might be different, the corrected group mean difference may be different for different values of the covariate (i.e., different practgrp simple effects for different values of numpract). So, it is usually a good idea to ask for group comparisons at several values of the covariate.

For this analysis, it makes sense to ask for group comparisons for 1, 3, 5, 7, & 9 practices. However, remember that the number of practices variable we've included in the model has been mean-centered. So, we have to take that mean centering into account!

- 1 raw practices corresponds with a mean-centered value of  $1 - 6.6458 = -5.6458$
- 3 raw practices corresponds with a mean-centered value of  $3 - 6.6458 = -3.6458$
- 5 raw practices corresponds with a mean-centered value of  $5 - 6.6458 = -1.6458$
- 7 raw practices corresponds with a mean-centered value of  $7 - 6.6458 = .3542$
- 9 raw practices corresponds with a mean-centered value of  $9 - 6.6458 = 2.3542$

```
UNIANOVA testperf BY practgrp WITH numpract_cen
/METHOD = SSTYPE(3)
/EMMEANS = TABLES(practgrp) WITH (numpract_cen = -5.6458) COMPARE (practgrp)
/EMMEANS = TABLES(practgrp) WITH (numpract_cen = -3.6458) COMPARE (practgrp)
/EMMEANS = TABLES(practgrp) WITH (numpract_cen = -1.6458) COMPARE (practgrp)
/EMMEANS = TABLES(practgrp) WITH (numpract_cen = .3542) COMPARE (practgrp)
/EMMEANS = TABLES(practgrp) WITH (numpract_cen = 2.3542) COMPARE (practgrp)
/PRINT = DESCRIPTIVE PARAMETER
/DESIGN=practgrp numpract_cen numpract_cen*practgrp.
```

## Full Model ANCOVA output

### Tests of Between-Subjects Effects

Dependent Variable: testperf

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	15045.388 <sup>a</sup>	5	3009.078	45.889	.000
Intercept	163361.724	1	163361.724	2491.273	.000
practgrp	9096.152	2	4548.076	69.358	.000
numpract_cen	154.926	1	154.926	2.363	.132
practgrp * numpract_cen	2709.648	2	1354.824	20.661	.000
Error	2754.091	42	65.574		
Total	248125.000	48			
Corrected Total	17799.479	47			

a. R Squared = .845 (Adjusted R Squared = .827)

### Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	85.940	2.548	33.724	.000
[practgrp=1.00]	-40.228	3.440	-11.695	.000
[practgrp=2.00]	-18.372	3.255	-5.645	.000
[practgrp=3.00]	0 <sup>a</sup>	.	.	.
numpract_cen	2.645	1.093	2.421	.020
[practgrp=1.00] * numpract_cen	-6.164	1.353	-4.557	.000
[practgrp=2.00] * numpract_cen	.647	1.356	.477	.636
[practgrp=3.00] * numpract_cen	0 <sup>a</sup>	.	.	.

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

The F table shows that we have a significant practgrp main effect after controlling for number of practices and the interaction,

There is not a significant main effect of numpract after controlling for practgrp and the interaction

There is a significant interaction of practgrp and numpract (which means one or both main effects might be misleading!)

Notice that the SSerror is much smaller in this ANCOVA model than in the original ANOVA model (5517.188).

For each t-test, df = 42, the dferror from the ANOVA model just above.

## Simple Effects of “practgrp”

Here are the five simple “EMMEANS” analyses. They represent the simple effect of practice group for each of five different amounts of practice (1, 3, 5, 7 & 9). Since the pairwise comparisons provide more information than the tests, I have presented just the former, to save space.

### 1 practice Easier < Same = Harder

#### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
easier	65.583 <sup>a</sup>	3.949
same	48.980 <sup>a</sup>	4.949
harder	71.007 <sup>a</sup>	7.977

a. Covariates appearing in the model are evaluated at the following values:  
numpract\_cen = -5.65.

#### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	16.602 <sup>*</sup>	6.331	.012
	harder	-5.424	8.901	.546
same	easier	-16.602 <sup>*</sup>	6.331	.012
	harder	-22.026 <sup>*</sup>	9.387	.024
harder	easier	5.424	8.901	.546
	same	22.026 <sup>*</sup>	9.387	.024

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

### 3 practices Easier = Same < Harder

#### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
easier	58.544 <sup>a</sup>	2.706
same	55.565 <sup>a</sup>	3.545
harder	76.297 <sup>a</sup>	5.890

a. Covariates appearing in the model are evaluated at the following values:  
numpract\_cen = -3.65.

#### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	2.979	4.460	.508
	harder	-17.753 <sup>*</sup>	6.482	.009
same	easier	-2.979	4.460	.508
	harder	-20.732 <sup>*</sup>	6.874	.004
harder	easier	17.753 <sup>*</sup>	6.482	.009
	same	20.732 <sup>*</sup>	6.874	.004

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

### 5 practices Easier < Same < Harder

#### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
easier	51.505 <sup>a</sup>	2.034
same	62.150 <sup>a</sup>	2.408
harder	81.587 <sup>a</sup>	3.911

a. Covariates appearing in the model are evaluated at the following values:  
numpract\_cen = -1.65.

#### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	-10.645 <sup>*</sup>	3.153	.002
	harder	-30.082 <sup>*</sup>	4.408	.000
same	easier	10.645 <sup>*</sup>	3.153	.002
	harder	-19.437 <sup>*</sup>	4.593	.000
harder	easier	30.082 <sup>*</sup>	4.408	.000
	same	19.437 <sup>*</sup>	4.593	.000

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

### 7 practices Easier < Same < Harder

#### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
easier	44.466 <sup>a</sup>	2.459
same	68.735 <sup>a</sup>	2.047
harder	86.877 <sup>a</sup>	2.334

a. Covariates appearing in the model are evaluated at the following values:  
numpract\_cen = .35.

#### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	-24.269 <sup>*</sup>	3.200	.000
	harder	-42.411 <sup>*</sup>	3.390	.000
same	easier	24.269 <sup>*</sup>	3.200	.000
	harder	-18.142 <sup>*</sup>	3.104	.000
harder	easier	42.411 <sup>*</sup>	3.390	.000
	same	18.142 <sup>*</sup>	3.104	.000

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

### 9 practices Easier < Same < Harder

#### Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
easier	37.427 <sup>a</sup>	3.613
same	75.319 <sup>a</sup>	2.781
harder	92.167 <sup>a</sup>	2.269

a. Covariates appearing in the model are evaluated at the following values:  
numpract\_cen = 2.35.

#### Pairwise Comparisons

Dependent Variable: testperf

(I) practgrp	(J) practgrp	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>
easier	same	-37.892 <sup>*</sup>	4.559	.000
	harder	-54.740 <sup>*</sup>	4.266	.000
same	easier	37.892 <sup>*</sup>	4.559	.000
	harder	-16.848 <sup>*</sup>	3.589	.000
harder	easier	54.740 <sup>*</sup>	4.266	.000
	same	16.848 <sup>*</sup>	3.589	.000

Based on estimated marginal means

\*. The mean difference is significant at the .050 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).



## Simple Effects of “numpract”

The simple effect of the quantitative variable for each IV group is represented as the slope of the covariate-DV regression line for that group.

The problem is that we only get part of the information we need to describe the interaction this way from the ANCOVA

- from the ANOVA table we get the F-test of the interaction, which tells us whether or not the slope of the covariate-DV regression line is significantly different for the three groups
- from the Parameter Estimates table, we get the model the covariate-DV regression line *for the group originally coded “3”* (Harder practice) and an t-test of whether the slope is significantly different from 0 (flat).
- **but**, we don’t get the model or significance test for the covariate-DV regression line *for either the group originally coded “1”* (Easier) or *the group originally coded “2”* (Similar).

To get the slope of the covariate-DV regression line for the group coded “1”, we have to recode the group variable, and then rerun the ANCOVA, using the recoded group variable.

Getting regression model for Easier group

recode practgrp (1=3) (2=2) (3=1) into practgrpSHE.

```
UNIANOVA testperf BY practgrpSHE WITH
numpract_cen
/METHOD = SSTYPE(3)
/PRINT = PARAMETER
/DESIGN=practgrpSHE numpract_cen
numpract_cen*practgrpSHE.
```

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	45.713	2.311	19.784	.000
[practgrpSHE=1.00]	40.228	3.440	11.695	.000
[practgrpSHE=2.00]	21.856	3.072	7.114	.000
[practgrpSHE=3.00]	0 <sup>a</sup>	.	.	.
numpract_cen	-3.519	.798	-4.411	.000
[practgrpSHE=1.00] * numpract_cen	6.164	1.353	4.557	.000
[practgrpSHE=2.00] * numpract_cen	6.812	1.132	6.018	.000
[practgrpSHE=3.00] * numpract_cen	0 <sup>a</sup>	.	.	.

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

Getting regression model for Similar group

recode practgrp (1=1) (2=3) (3=2) into practgrpEHS.

```
UNIANOVA testperf BY practgrpEHS WITH
numpract_cen
/METHOD = SSTYPE(3)
/PRINT = PARAMETER
/DESIGN=practgrpEHS numpract_cen
numpract_cen*practgrpEHS.
```

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	67.568	2.025	33.375	.000
[practgrpEHS=1.00]	-21.856	3.072	-7.114	.000
[practgrpEHS=2.00]	18.372	3.255	5.645	.000
[practgrpEHS=3.00]	0 <sup>a</sup>	.	.	.
numpract_cen	3.292	.803	4.101	.000
[practgrpEHS=1.00] * numpract_cen	-6.812	1.132	-6.018	.000
[practgrpEHS=2.00] * numpract_cen	-.647	1.356	-.477	.636
[practgrpEHS=3.00] * numpract_cen	0 <sup>a</sup>	.	.	.

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

The models relating Number of Practices to Test Performance for the three groups are:

Harder group  $\text{testperf}' = ( 2.645 * \text{numpract\_cen} ) + 85.940$  the slope is significantly positive  $p = .02$

Similar group  $\text{testperf}' = ( 3.292 * \text{numpract\_cen} ) + 67.568$  the slope is significantly positive  $p < .001$

Easier group  $\text{testperf}' = ( -3.519 * \text{numpract\_cen} ) + 45.713$  the slope is significantly negative  $p < .001$

## Plotting the Full Model ANCOVA Results

You can use the Parameter Estimates from either ANCOVA we just did – which group is coded “1” and which is coded “2” doesn’t change the overall model, just how it is expressed in the regression weight. This example will use the parameters from the initial analysis with easier practice coded as “2”. Remember to:

- Use the IV group originally coded as “1” as the “z1 wt = 1 & z2 wt = 0” group (Easier)
- Use the IV group originally coded as “2” as the “z1 wt = 0 & z2 wt =1” group (Similar)
- Use the IV group originally coded as “3” as the “z1 wt = 0 & z2 wt =0” group (Harder)

### Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	85.940	2.548	33.724	.000
[practgrp=1.00]	-40.228	3.440	-11.695	.000
[practgrp=2.00]	-18.372	3.255	-5.645	.000
[practgrp=3.00]	0 <sup>a</sup>	.	.	.
numpract_cen	2.645	1.093	2.421	.020
[practgrp=1.00] * numpract_cen	-6.164	1.353	-4.557	.000
[practgrp=2.00] * numpract_cen	.647	1.356	.477	.636
[practgrp=3.00] * numpract_cen	0 <sup>a</sup>	.	.	.

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

<b>Constant</b>	constant	84.94				
<b>Covariate</b>	b(x)	2.645		Practice	z1 wt	Z2 wt
<b>IV/Groups - 1st</b>	b(z1)	-40.228		Harder	0	0
<b>Interaction -- 1st</b>	b(xz1)	-6.164		Easier	1	0
<b>IV/Groups - 2nd</b>	b(z2)	-18.372		Similar	0	1
<b>Interaction -- 2nd</b>	b(xz2)	0.647				
	x(mean)	6.6458				
	x(std)	2.6143				

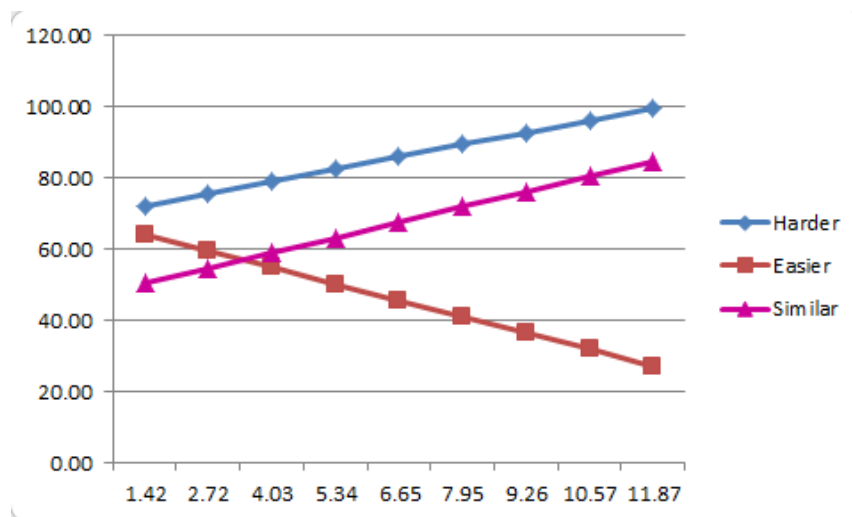


Figure 1:  
Plot of the ANCOVA model

Write-up for the ANCOVA

An ANCOVA was performed including Practice Difficulty Group (Easier, Similar & Harder Difficulty), Number of Practices and their interaction. The plot of the ANCOVA model is shown in Figure 1.

There is an interaction of Practice Item Difficulty and Number of Practices as they relate to Test Performance,  $F(2, 42) = 20.661$ ,  $MSe = 65.574$ ,  $p < .001$ . The pattern of the interaction is that, as can be seen in Figure 1, the Harder and Easier groups performed significantly better than the Similar group following 1 practice, the Harder group perform significantly better than the Same and Easier groups following 3 practices, and following 5, 7 & 9 practices, the Harder group performed best while the Easier group performed poorest.

An alternative description of the pattern of the interaction is that the slope of the Number of Practice regression line is positive for the Harder,  $b = 2.645$ ,  $p = .02$ , and Similar,  $b = 3.292$ ,  $p < .001$ , groups, while this slope is negative for the Easier group,  $b = -3.519$ ,  $p < .001$ .

The main effect for Number of Practices was non-significant,  $F(1,42) = 2.363$ ,  $MSe = 65.574$ ,  $p = .132$ . However this main effect was not descriptive for any of the three Practice Difficulty groups, because of the pattern of the interaction. Although there is no relationship between number of practices and test performance *on average*, there was a positive relationship for the Harder and Similar groups and a negative relationship for Easier group.

The main effect of Practice Group was significant,  $F(1,42) = 69358$ ,  $MSe = 65.574$ ,  $p < .001$ . However, this main effect was not descriptive, as whether the pattern of Test Performance differences among the groups changed with the Number of Practices.