2-Group ANCOVA

The purpose of the study was to compare the Test Performance of: 1) students who had prepared for the test using practice problems that were similar in difficulty to the actual test problems (Same) and 2) students who had prepared using practice problems that were easier than the actual test problems (Easier). Students were randomly assigned to one of the practice groups and given a packet that had instructions for the problems, 12 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	same	16	77.5000	12.90994
	easier	16	60.6250	12.89380
	Total	32	69.0625	15.31589
numpract	same	16	6.6250	2.60448
	easier	16	5.2500	2.62043
	Total	32	5.9375	2.66322

Those who practice with same difficulty problems performed significantly better than those who practiced with the easier problems.

While there is not a statistically significant difference between the number or practices completed by the group, the difference is about ½ std, which is probably too large to treat as "equivalent"

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
testperf	Between Groups	2278.125	1	2278.125	13.686	.001
	Within Groups	4993.750	30	166.458		
	Total	7271.875	31			
numpract	Between Groups	15.125	1	15.125	2.216	.147
	Within Groups	204.750	30	6.825		
	Total	219.875	31			

This would seem to be an "augmenting confound" because the group expected to perform better had the higher mean on the potential confound.

Correlations

		testperf	numpract
testperf	Pearson Correlation	1	.325
	Sig. (2-tailed)		.061
	N	32	32
numpract	Pearson Correlation	.325	1
	Sig. (2-tailed)	.061	
	N	32	32

Similarly, while there is not a significant correlation between number of practice and test performance, the correlation is "middle sized" and too large to be treated as "unrelated to the DV".

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 - 5.9375.
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 both 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 and Similar 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 differnt 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
same	77.5000	12.90994	16
easier	60.6250	12.89380	16
Total	69.0625	15.31589	32

Tests of Between-Subjects Effects

Dependent Variable: testperf

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2281.819[a]	2	1140.909	6.630	.004
Intercept	152628.125	1	152628.125	887.007	.000
practgrp	2168.102	1	2168.102	12.600	.001
numpract_cen	3.694	1	3.694	.021	.885
Error	4990.056	29	172.071		
Total	159900.000	32			
Corrected Total	7271.875	31			

a. R Squared = .314 (Adjusted R Squared = .266)

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	60.533	3.339	18.127	.000	53.703	67.363
[practgrp=1.00]	17.060	4.806	3.550	.001	7.230	26.889
[practgrp=2.00]	0[a]	.	.	.	.	.
numpract_cen	-.134	.917	-.147	.885	-2.009	1.741

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

These are the same (uncorrected) 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 both groups.

Main Effects Model Corrected Means and their Comparison

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
same	77.592[a]	3.339	70.762	84.422
easier	60.533[a]	3.339	53.703	67.363

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.[b]
same	easier	17.060*	4.806	.001
easier	same	-17.060*	4.806	.001

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	2168.102	1	2168.102	12.600	.001
Error	4990.056	29	172.071		

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 77.592 for the Same Difficulty group and 60.533 for the Easier group.

The F-table and pairwise comparions (which are the same when comparing 2 groups) tell us that this 17.06 mean difference is statistically significant.

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 "2xQ 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 interaction
- 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.

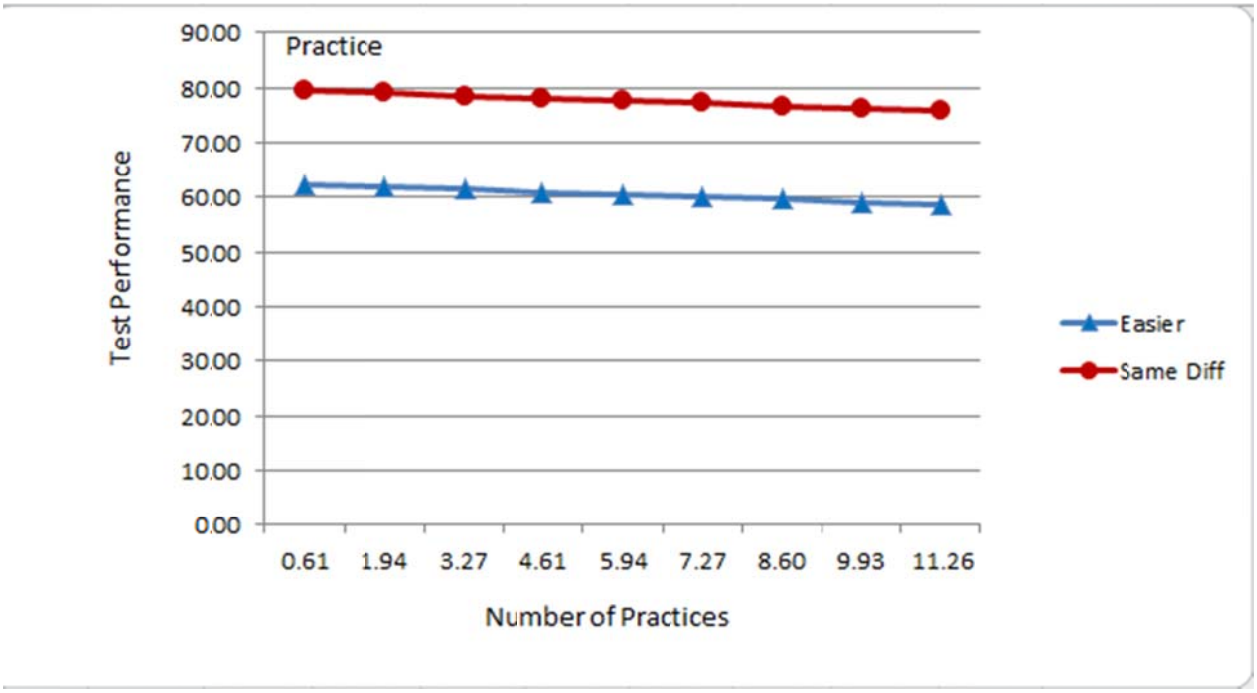
Intercept	constant	60.533		
Covariate	b(x)	-0.352	Practice Group	*z wt*
IV/Groups	b(z)	17.06	Easier	*0*
Interaction	b(xz)	0	Same Diff	*1*
Covariate	x(mean)	5.9375		
	x(std)	2.66322		

Please note:

Use the IV group originally coded as "1" as the "Z wt = 1" group (same difficulty in this example)

Use the IV group originally coded as "2" as the "Z wt = 0" group (easier in this example).

The program makes a pretty decent plot of the results…



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

- 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, 6, 9 & 12 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 – 5.9375 = -4.9375
- 3 raw practices corresponds with a mean-centered value of 3 – 5.9375 = -2.9375
- 6 raw practices corresponds with a mean-centered value of 6 – 5.9375 = .0625
- 9 raw practices corresponds with a mean-centered value of 9 – 5.9375 = 3.0625
- 12 raw practices corresponds with a mean-centered value of 12 – 5.9375 = 6.0625

```
UNIANOVA testperf BY practgrp WITH numpract_cen
  /METHOD = SSTYPE(3)
  /EMMEANS = TABLES(practgrp) WITH (numpract_cen = -4.9375) COMPARE (practgrp)
  /EMMEANS = TABLES(practgrp) WITH (numpract_cen = -2.9375) COMPARE (practgrp)
  /EMMEANS = TABLES(practgrp) WITH (numpract_cen = .0625) COMPARE (practgrp)
  /EMMEANS = TABLES(practgrp) WITH (numpract_cen = 3.0625) COMPARE (practgrp)
  /EMMEANS = TABLES(practgrp) WITH (numpract_cen = 6.0625) 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	4656.862[a]	3	1552.287	16.621	.000
Intercept	132654.284	1	132654.284	1420.383	.000
practgrp	2160.835	1	2160.835	23.137	.000
numpract_cen	2.638	1	2.638	.028	.868
practgrp * numpract_cen	2375.044	1	2375.044	25.431	.000
Error	2615.013	28	93.393		
Total	159900.000	32			
Corrected Total	7271.875	31			

a. R Squared = .640 (Adjusted R Squared = .602)

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	58.205	2.503	23.253	.000
[practgrp=1.00]	17.031	3.541	4.810	.000
[practgrp=2.00]	0[a]	.	.	.
numpract_cen	-3.519	.952	-3.696	.001
[practgrp=1.00] * numpract_cen	6.812	1.351	5.043	.000
[practgrp=2.00] * numpract_cen	0[a]	.	.	.

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 (4990.056).

df for each t-test = dferror = 28

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, 6, 9 & 12). Since the pairwise comparisons are redundant with the univariate tests, I have presented just the latter, to save space.

1 practice

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
same	58.980[a]	5.906
easier	75.583[a]	4.713

a. Covariates appearing in the model are evaluated at the following values: numpract_cen = -4.94.

Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	450.874	1	450.874	4.828	.036
Error	2615.013	28	93.393		

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

3 practices

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
same	65.565[a]	4.231
easier	68.544[a]	3.229

a. Covariates appearing in the model are evaluated at the following values: numpract_cen = -2.94.

Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	29.252	1	29.252	.313	.580
Error	2615.013	28	93.393		

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

6 practices

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
same	75.442[a]	2.489
easier	57.985[a]	2.519

a. Covariates appearing in the model are evaluated at the following values: numpract_cen = .06.

Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	2269.100	1	2269.100	24.296	.000
Error	2615.013	28	93.393		

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

9 practices

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
same	85.319[a]	3.319
easier	47.427[a]	4.311

a. Covariates appearing in the model are evaluated at the following values: numpract_cen = 3.06.

Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	4529.900	1	4529.900	48.503	.000
Error	2615.013	28	93.393		

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

12 practices

Estimates

Dependent Variable: testperf

practgrp	Mean	Std. Error
same	95.197[a]	5.688
easier	36.869[a]	6.867

a. Covariates appearing in the model are evaluated at the following values: numpract_cen = 6.06.

Univariate Tests

Dependent Variable: testperf

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	3996.412	1	3996.412	42.791	.000
Error	2615.013	28	93.393		

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

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 two groups
- from the Parameter Estimates table, we get the model the covariate-DV regression line *for the group originally coded "2"* (easier practice) and an t-test of whether the slope is significantly different from 0 (flat).
    - for the easier practice group (coded "2") testperf' = (-3.519 * numpract_cen) + 58.205
    - this regression slope is significantly negative, t(28) = -3.696, p < .001
- **but**, we don't get the model of the covariate-DV regression line *for the group originally coded "1"* (similar practice) or a t-test of whether the slope is significantly different from 0 (flat).

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

```
recode practgrp (1=2) (2=1) into practgrp21.

UNIANOVA testperf BY practgrp21 WITH numpract_cen
  /METHOD = SSTYPE(3)
  /PRINT = PARAMETER
  /DESIGN=practgrp21 numpract_cen numpract_cen*practgrp21.
```

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	75.236	2.504	30.044	.000	70.107	80.366
[practgrp21=1.00]	-17.031	3.541	-4.810	.000	-24.284	-9.778
[practgrp21=2.00]	0[a]	.	.	.	.	.
numpract_cen	3.292	.958	3.437	.002	1.330	5.255
[practgrp21=1.00] * numpract_cen	-6.812	1.351	-5.043	.000	-9.579	-4.045
[practgrp21=2.00] * numpract_cen	0[a]	.	.	.	.	.

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

From this Parameter Estimates table, we get the model the covariate-DV regression line *for the group originally coded "1"* but now coded "2" (similar difficulty practice) and an t-test of whether the slope is significantly different from 0 (flat).

- for the similar practice group (now coded "2") testperf' = (3.292 * numpract_cen) + 75.236
- this regression slope is significantly positive, t= 3.437, p = .002

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 "Z wt = 1" group (Same Difficulty in this example)
- use the IV group originally coded as "2" as the "Z wt = 0" group (Easier in this example).

Parameter Estimates

Dependent Variable: testperf

Parameter	B	Std. Error	t	Sig.
Intercept	58.205	2.503	23.253	.000
[practgrp=1.00]	17.031	3.541	4.810	.000
[practgrp=2.00]	0[a]	.	.	.
numpract_cen	-3.519	.952	-3.696	.001
[practgrp=1.00] * numpract_cen	6.812	1.351	5.043	.000
[practgrp=2.00] * numpract_cen	0[a]	.	.	.

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

Intercept	constant	58.205		
Covariate	b(x)	-3.519	**Practice Group**	***z wt***
IV/Groups	b(z)	17.031	Easier	***0***
Interaction	b(xz)	6.812	Same Diff	***1***
Covariate	x(mean)	5.9375		
	x(std)	2.66322		

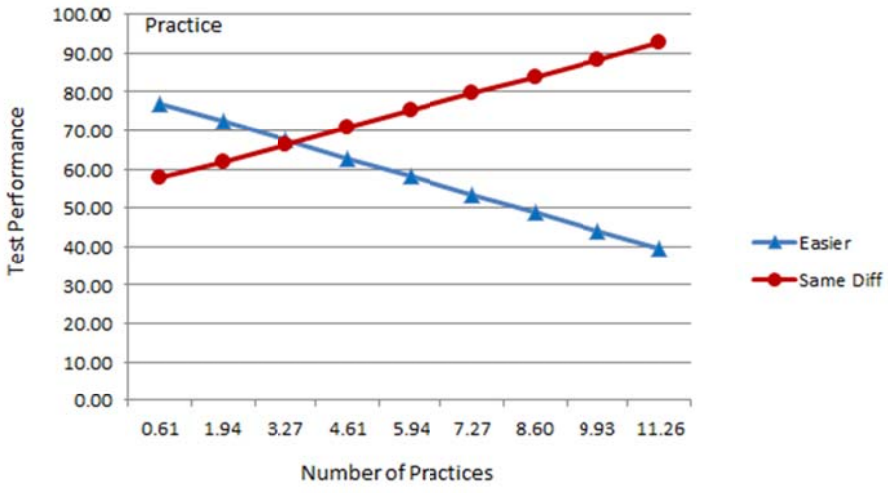


Figure 1:
Plot of the ANCOVA model

Write-up for the ANCOVA

An ANCOVA was performed including Practice Difficulty Group (Easier & Similar 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, $F(1, 28) = 25.431$, $MSe = 93.393$, $p < .001$. The pattern of the interaction is that, as can be seen in Figure 1, the Easier Practice group performed significantly better than the Similar Difficulty group following 1 practice ($p=.034$), there was no significant difference following 3 practices ($p = .580$), while the Similarly Difficult group performed significantly better than the Easier Practice group following 6, 9 & 12 practices ($p < .001$ for each).

An alternative description of the pattern of the interaction is that the slope of the Number of Practice regression line is positive for the Similar Difficulty group, $b = 3.292$, $p = .002$, while this slope is negative for the Easier group, $b = -3.519$, $p < .001$

The main effect for Number of Practices was non-significant, $F(1,28)=.028$, $MSe = 93.393$, $p = .868$. However this main effect was not descriptive for either Practice Difficulty group, 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 Same Difficulty practices and a negative relationship for Easier practices.

The main effect of Practice Group was significant, $F(1,28) = 23.137$, $MSe = 93.393$, $p < .001$. However, this main effect was not descriptive, as whether the Easier or Similar Difficulty practice group performed was different for different Number of Practices.