

# 10. Comparing Means Using Repeated Measures ANOVA

## Objectives

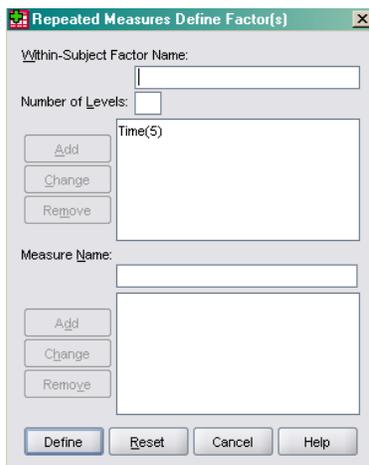
- ◆ Calculate repeated measures ANOVAs
- ◆ Calculate effect size
- ◆ Conduct multiple comparisons
- ◆ Graphically illustrate mean differences

Repeated measures ANOVAs are used to examine mean differences in related variables. Typically the independent variable is either time (e.g., depression is measured in the same group of people at multiple points in time) or condition (e.g., each subject receives every condition). In SPSS, we will use the General Linear Model to calculate repeated measures ANOVAs.

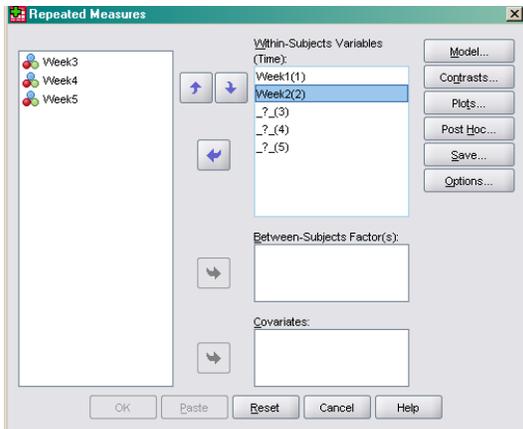
## Using GLM Repeated Measures to Calculate Repeated Measures ANOVAs

Let's begin with an example from the exercises in Chapter 18 in the *Fundamentals* book. In this example, the duration of migraine headaches was recorded among the same group of individuals over 5 weeks. The first 2 weeks were part of a baseline period and the final 3 weeks were part of an intervention period in which subjects were trained to apply relaxation techniques. In this case, the independent variable is time and the dependent variable is headache duration.

- ✓ **Open** *migraines.sav*.
- ✓ Select **Analyze/General Linear Model/Repeated Measures**.

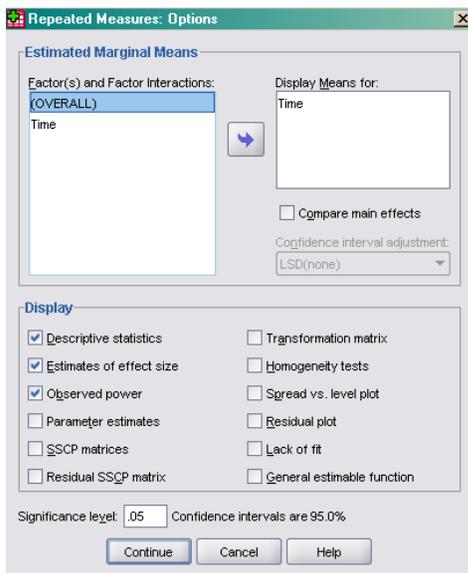
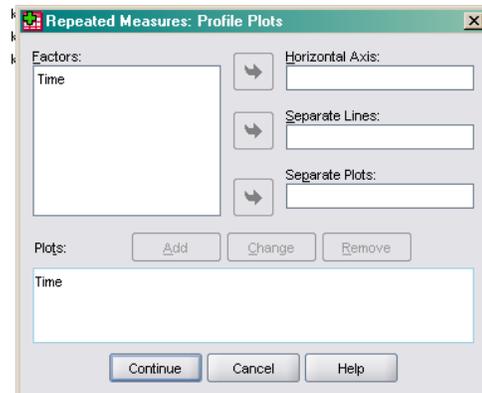


- ✓ The default for **Within-Subject Factor Name** is Factor 1. Let's change it to Time by typing in the box. Specify 5 for **Number of Levels** since there are 5 weeks and then click **Add**. Next, click **Define**.



✓ As you can see, there are 5 spots under **Within-Subject Variables**. We need to indicate that each week is a variable, and we want to list them in the right order. You can either select them one at a time and arrow them into the variable list, or you can select them all by holding down the control button while you select each one, and arrow them in at once. This is a simple design, so there are no Between-Subject Factors or Covariates. Click **Plots**.

✓ Select time as the **Horizontal Axis** or Y Axis. Then click **Add**, and **Continue**. This will plot the mean duration of headaches for each week for us. In the main dialog box, click **Options**.



✓ Under **Display**, select **Descriptive Statistics**, **Estimates of Effect Size**, and **Observed Power**. Then click **Continue**. And finally, **Ok**. The output follows.

**Descriptive Statistics**

	Mean	Std. Deviation	N
headache duration week 1	20.78	7.17	9
headache duration week 2	20.00	10.22	9
headache duration week 3	9.00	3.12	9
headache duration week 4	5.78	3.42	9
headache duration week 5	6.78	4.12	9

**Multivariate Tests<sup>c</sup>**

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Observed Power <sup>b</sup>
Time Pillai's Trace	.909	12.548 <sup>a</sup>	4.000	5.000	.008	.909	.952
Wilks' Lambda	.091	12.548 <sup>a</sup>	4.000	5.000	.008	.909	.952
Hotelling's Trace	10.038	12.548 <sup>a</sup>	4.000	5.000	.008	.909	.952
Roy's Largest Root	10.038	12.548 <sup>a</sup>	4.000	5.000	.008	.909	.952

a. Exact statistic

b. Computed using alpha = .05

c. Design: Intercept  
Within Subjects Design: Time

**Mauchly's Test of Sphericity<sup>b</sup>**

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	.030	22.516	9	.009	.422	.522	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept  
Within Subjects Design: Time

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Time	Sphericity Assumed	1934.533	4	483.633	21.463	.000	85.852	1.000
	Greenhouse-Geisser	1934.533	1.687	1146.534	21.463	.000	36.214	.999
	Huynh-Feldt	1934.533	2.089	926.174	21.463	.000	44.831	1.000
	Lower-bound	1934.533	1.000	1934.533	21.463	.002	21.463	.981
Error(Time)	Sphericity Assumed	721.067	32	22.533				
	Greenhouse-Geisser	721.067	13.498	53.419				
	Huynh-Feldt	721.067	16.710	43.152				
	Lower-bound	721.067	8.000	90.133				

a. Computed using alpha = .05

**Tests of Within-Subjects Contrasts**

Measure: MEASURE\_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Time	Linear	1604.444	1	1604.444	36.181	.000	.819	36.181	.999
	Quadratic	82.571	1	82.571	6.215	.037	.437	6.215	.591
	Cubic	187.778	1	187.778	17.066	.003	.681	17.066	.950
	Order 4	59.740	1	59.740	2.779	.134	.258	2.779	.312
Error(Time)	Linear	354.756	8	44.344					
	Quadratic	106.286	8	13.286					
	Cubic	88.022	8	11.003					
	Order 4	172.003	8	21.500					

a. Computed using alpha = .05

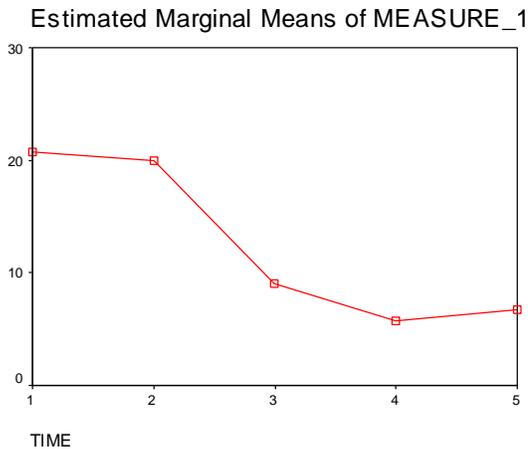
### Tests of Between-Subjects Effects

Measure: MEASURE\_1  
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power <sup>a</sup>
Intercept	6993.800	1	6993.800	67.119	.000	.894	1.000
Error	833.600	8	104.200				

a. Computed using alpha = .05

### Profile Plots



As you can see, there is a lot of output, much of which we can ignore for our purposes. Specifically, ignore Multivariate Tests, Tests of Within-Subjects Contrasts, and Tests of Between Subjects Effects. The multivariate tests are another way to run the analysis, and often not a good way. The contrasts give tests of linear and quadratic trends in the data, and are not particularly of interest here. There are no between subjects factors, so that output is not of interest. Now, let's look at the rest and compare it to the answers in the text. First, you can compare the mean scores for each week by looking at the **Descriptive Statistics** table. The next piece is **Mauchly's Test of Sphericity**, which tests the assumption that each of the time periods is approximately equally correlated with every other score. As noted in the text, when this assumption is violated, various corrections are applied. Also, as noted in the text, this is not a particularly good test, but it is about the best we have. The next table of interest, **Tests of Within-Subjects Effects**, is what we really want to see. Compare the textbook values to those listed in the rows marked Sphericity Assumed, because they were calculated the same way. As you can see, they are in agreement.

Now, note the values for eta squared and observed power. Can you interpret them? Nearly 73% of the variability in headache duration is accounted for by time. Observed power is based on the assumption that the true difference in population means is the difference implied by the sample means. Typically, we want to calculate power going into an experiment based on anticipated or previous effect size in other similar studies. This is useful in making decisions about sample size. So, observed power calculated here is not particularly useful.

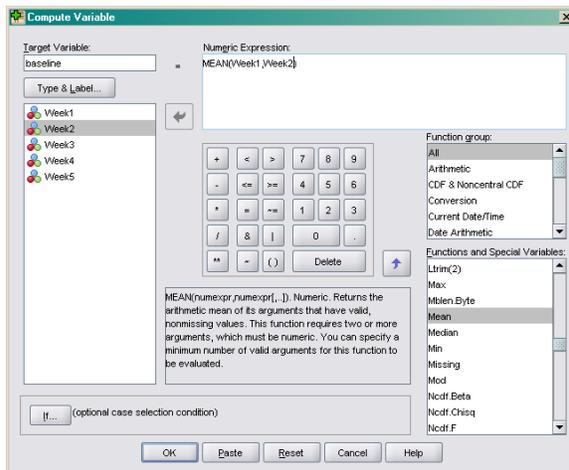
The graph is a nice illustration of the mean headache duration over time. You may want to edit it to include more meaningful labels and a title.

Now, we need to calculate multiple comparisons to help us understand the meaning of the significant effect of time on headache duration.

## Multiple Comparisons

Let's just try one of the possible multiple comparisons, the comparison between the overall baseline mean and the overall training mean. We can use SPSS Transform/Compute to calculate these averages for us rather than doing it manually.

- ✓ In the Data Editor window, select **Transform/Compute**.



- ✓ Type baseline under **Target Variable**. Under the list of **Functions**, select **MEAN** and arrow it into the dialog box. We need to tell SPSS from what variables to calculate the mean. Select week1 and week2 to replace the 2 question marks. Make sure they are separated by a comma and the question marks are gone. Then, click **Ok**.

- ✓ Look at the new variable in the Data Editor. Does it look right?
- ✓ Click **Transform/Compute** again. Click **Reset** to remove the previous information. Name the next **Target Variable** training. Select **MEAN** again. Specify, week3, week4, and week5. Make sure the question marks are gone and commas separate each variable. Then, click **Ok**. Check out your new variable.
- ✓ Use **Analyze/Descriptives** to calculate the means for baseline and training. The data follow.

### Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
baseline	9	6.00	31.50	20.3889	8.31331
Training	9	4.00	13.00	7.1852	2.84366
Valid N (listwise)	9				

As you can see, the means are consistent with those reported in the textbook. Now, you can apply formula using  $MS_{\text{error}}$  from the ANOVA. The computations follow.

$$t = \frac{20.39 - 7.17}{\sqrt{22.53\left(\frac{1}{18} + \frac{1}{27}\right)}} = \frac{13.20}{\sqrt{2.086}} = 9.14$$

Although some hand calculations are required, we saved time and reduced the likelihood of making errors by using SPSS to compute the new mean scores for baseline and training for us.

In this chapter, you learned to use the General Linear Model to calculate repeated measures ANOVAs. In addition, you learned to use SPSS to calculate new means for use in multiple comparisons. Try the following exercises to help you become more familiar with the process.

### Exercises

The following exercises are based on *Eysenck repeated.sav*.

1. Use a repeated measures ANOVA to examine the effect of condition on recall. Compare your results to those presented in the textbook in Section 18.7.
2. Use SPSS to calculate the effect size of condition.
3. Plot the mean difference in recall by conditions.
4. Use SPSS to calculate the mean of counting, rhyming, adjective, and intentional and label it lowproc for lower processing. Then use the multiple comparisons procedure explained in the textbook to compare the mean recall from the lower processing conditions to the mean recall for imagery, which was the highest processing condition. Write a brief statement explaining the results.