



## Contrasts and Post Hoc Tests for One-Way Independent ANOVA Using SPSS

### Running the Analysis

In last week's lecture we came across an example, from Field (2013), about the drug Viagra, which is a sexual stimulant (used to treat impotence), which is supposed to make men better lovers. Suppose we tested this belief by taking three groups of people and administering one group with a placebo (such as a sugar pill), one group with a low dose of Viagra and one with a high dose. The dependent variable was an objective measure of libido (I will tell you only that it was measured over the course of a week — the rest I shall leave to your own imagination). The data can be found below.

Table 1: Data showing how libido differs after different doses of Viagra

Dose	Libido
Placebo	3
	2
	1
	1
	4
Low Dose Viagra	5
	2
	4
	2
	3
High Dose Viagra	7
	4
	5
	3
	6

First, let's conduct an ANOVA on the Viagra data. As with the data from earlier in this lesson (rugby injuries) we need to enter the data into the data editor using a coding variable specifying to which of the three groups the score belongs. So, the data are entered in two columns (one called **dose** which specifies how much Viagra the person was given and one called **libido** which indicates the person's libido over the following week). You can code the variable **dose** any way you wish but I recommend something simple such as 1 = placebo, 2 = low dose and 3 = high dose.



- Enter the data into SPSS.
- Save the data onto a disk in a file called **Viagra.sav**.
- Plot an error bar graph of the data.

To conduct one-way ANOVA we have to first access the main dialogue box by selecting **Analyze** **Compare Means** **One-Way ANOVA...** (Figure 1). This dialogue box has a space where you can list one or more dependent variables and a second space to specify a grouping variable, or *factor*. Factor is another term for independent variable. For the Viagra data we need select only **libido** from the variable list and drag it to the box labelled *Dependent List* (or click on click on ). Then select the grouping variable **dose** and transfer it to the box labelled *Factor* (or click on ).

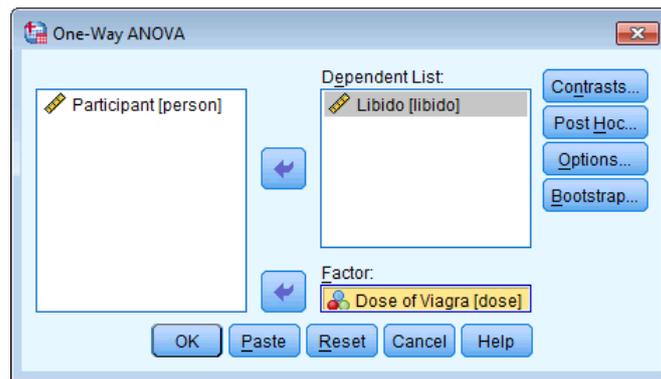


Figure 1: Dialogue box for one-way ANOVA

## Planned Comparisons Using SPSS

Click on **Contrasts...** to access the dialog box that allows you to conduct planned comparisons. The dialog box is shown in Figure 2 and has two sections. The first section is for specifying trend analyses. If you want to test for trends in the data then tick the box labelled **Polynomial** and select the degree of polynomial you would like. The Viagra data has only three groups and so the highest degree of trend there can be is a quadratic trend (see Field, 2013 Chapter 11). It is important from the point of view of trend analysis that we have coded the grouping variable in a meaningful order. We expect libido to be smallest in the placebo group, to increase in the low dose group and then to increase again in the high dose group. To detect a meaningful trend, we need to have coded these groups in ascending order. We have done this by coding the placebo group with the lowest value 1, the low dose group with the middle value 2, and the high dose group with the highest coding value of 3. If we coded the groups differently, this would influence both whether a trend is detected, and if by chance a trend is detected whether it is meaningful. For the Viagra data there are only three groups and so we should select the polynomial option ( **Polynomial**), and then select a quadratic degree by clicking on **Linear** and then selecting **Quadratic** (the drop-down list should now say **Quadratic**). If a quadratic trend is selected SPSS will test for both linear and quadratic trends.

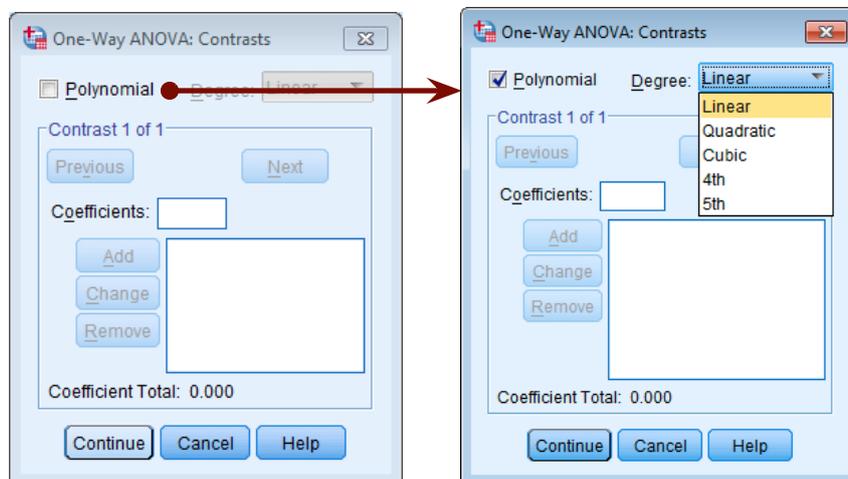


Figure 2: Dialog box for conducting planned comparisons

The lower part of the dialogue box in Figure 2 is for specifying planned comparisons. To conduct planned comparisons we need to tell SPSS what weightings to assign to each group. The first step is to decide which comparisons you want to do and then what weights must be assigned to each group for each of the contrasts (see Field, 2013, Chapter 11). We saw in the lecture that a sensible set of contrasts would be to compare the two experimental groups to the control group (Low dose + high dose vs. Placebo) as contrast 1, and then compare the low dose to the high dose in a second contrast. The weights for contrast 1 would be: -2 (placebo group), +1 (Low dose group), and +1 (high dose group). We will specify this contrast first. It is important to make sure that you enter the correct weighting for each group, so you should remember that the first weight that you enter should be the weight for the *first* group (that is, the group coded with the lowest value in the data editor). For the Viagra data, the group coded with the lowest value was the placebo



group (which had a code of 1) and so we should enter the weighting for this group first. Click in the box labelled *Coefficients* with the mouse and then type '-2' in this box and click on **Add**. Next, we input the weight for the second group, which for the Viagra data is the low dose group (because this group was coded in the data editor with the second highest value). Click in the box labelled *Coefficients* with the mouse and then type '1' in this box and click on **Add**. Finally, we need to input the weight for the last group, which for the Viagra data is the high dose group (because this group was coded with the highest value in the data editor). Click in the box labelled *Coefficients* with the mouse and then type '1' in this box and click on **Add**. The box should now look like **Figure 3** (left).

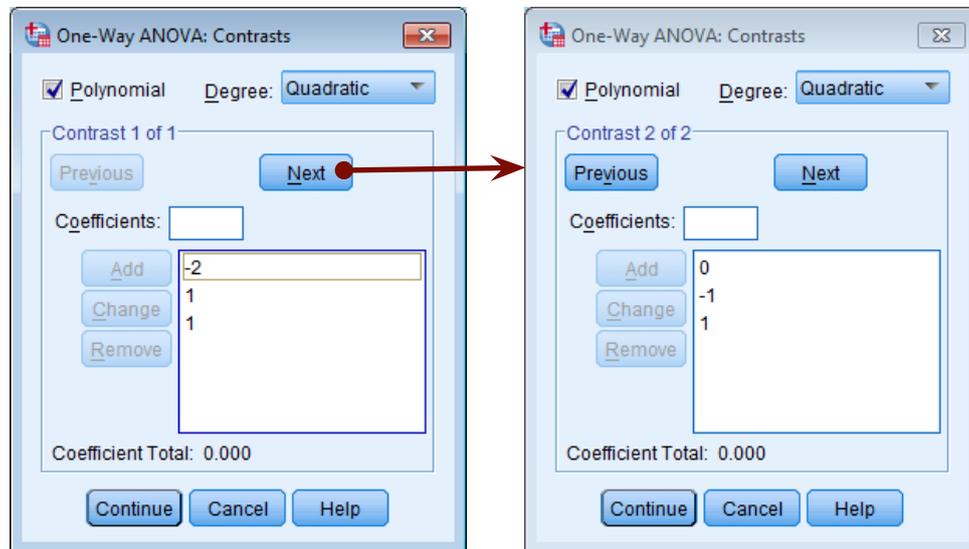


Figure 3: *Contrasts* dialog box completed for the two contrasts of the Viagra data

Once you have input the weightings you can change or remove any one of them by using the mouse to select the weight that you want to change. The weight will then appear in the box labelled *Coefficients* where you can type in a new weight and then click on **Change**. Alternatively, you can click on any of the weights and remove it completely by clicking **Remove**. Underneath the weights SPSS calculates the coefficient total, should equal zero (If you've used the correct weights). If the coefficient number is anything other than zero you should go back and check that the contrasts you have planned make sense and that you have followed the appropriate rules for assigning weights.

Once you have specified the first contrast, click on **Next**. The weightings that you have just entered will disappear and the dialogue box will now read *contrast 2 of 2*. The weights for contrast 2 should be: 0 (placebo group), +1 (Low dose group), and -1 (high dose group). We can specify this contrast as before. Remembering that the first weight we enter will be for the placebo group, we must enter the value zero as the first weight. Click in the box labelled *Coefficients* with the mouse and then type '0' and click on **Add**. Next, we need to input the weight for the low dose by clicking in the box labelled *Coefficients* and then typing '1' and clicking on **Add**. Finally, we need to input the weight for the high dose group by clicking in the box labelled *Coefficients* and then typing '-1' and clicking on **Add** (see Figure 3, right).

You should notice that the weights add up to zero as they did for contrast 1. It is imperative that you remember to input zero weights for any groups that are not in the contrast. When all of the planned contrasts have been specified click on **Continue** to return to the main dialogue box.

## Post Hoc Tests in SPSS

Once we have told SPSS which planned comparisons we have done, we can choose to do some post hoc tests. Normally if we have done planned comparisons we should not do post hoc tests (because we have already tested the hypotheses of interest). Likewise, if we choose to conduct post hoc tests then planned contrasts are unnecessary (because we have no hypotheses to test). However, for the sake of space we will conduct some post hoc tests on the Viagra data. Click on **Post Hoc...** in the main dialogue box to access the post hoc tests dialogue box (Figure 4). I recommend various post hoc procedures for various situations. The choice of comparison procedure will depend on the exact situation you have and whether it is more important for you to keep strict control over the familywise error rate or to have greater statistical power. I have drawn some general guidelines:



Field (2013) recommends:

- When you have equal sample sizes and you are confident that your population variances are similar then use R-E-G-W-Q or Tukey because both have good power and tight control over the Type I error rate.
- If sample sizes are slightly different then use Gabriel's procedure because it has greater power, but if sample sizes are very different use Hochberg's GT2.
- If there is any doubt that the population variances are equal then use the Games-Howell procedure because this seems to generally offer the best performance.

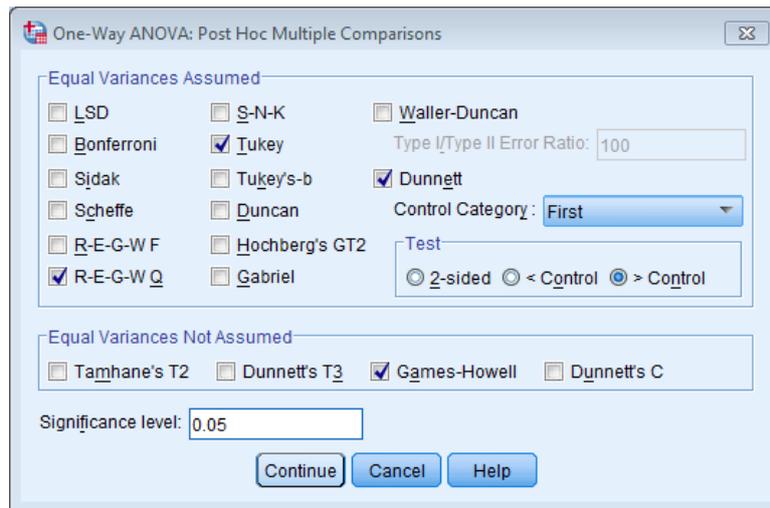


Figure 4: Dialogue box for specifying post hoc tests

I recommend running the Games-Howell procedure in addition to any other tests you might select because of the uncertainty of knowing whether the population variances are equivalent. For the Viagra data there are equal sample sizes and so we need not use Gabriel's test. We should use Tukey's test and R-E-G-W-Q and check the findings with the Games-Howell procedure. We have a specific hypothesis that both the high and low dose groups should differ from the placebo group and so we could use Dunnett's test to examine these hypotheses. Once you have selected Dunnett's test, change the control category (the default is to use the **Last** category) to specify that the **First** category be used as the control category (because the placebo group was coded with the lowest value). You can also choose whether to conduct a two-tailed test (**2-sided**), or a one-tailed test. If you choose a one-tailed test then you must predict whether you believe that the mean of the control group will be less than the experimental groups (**< Control**) or greater than the experimental groups (**> Control**). Because we predict that Viagra groups will have larger means than the control, chose a one-sided test (**> Control**). These are all of the post hoc tests that need to be specified and when the completed dialogue box looks like **Error! Reference source not found.** click on **Continue** to return to the main dialogue box.

## Options

The options for one-way ANOVA are fairly straightforward. First you can ask for some descriptive statistics, which will display a table of the means, standard deviations, standard errors, ranges and confidence intervals for the means of each group. This is a useful option to select because it assists in interpreting the final results. Select homogeneity of variance tests too (see last week's handout and your handout on bias).

We saw in the last handout that the main dialog box is the alluring **Bootstrap...** button. We aklos saw that selecting bootstrapping won't do anything useful to the main *F*-test; however, It will bootstrap contrasts and post hoc tests. For this example, we have a very small data set so bootstrapping is going to go haywire anyway, so we won't select it. Click on **OK** in the main dialog box to run the analysis.

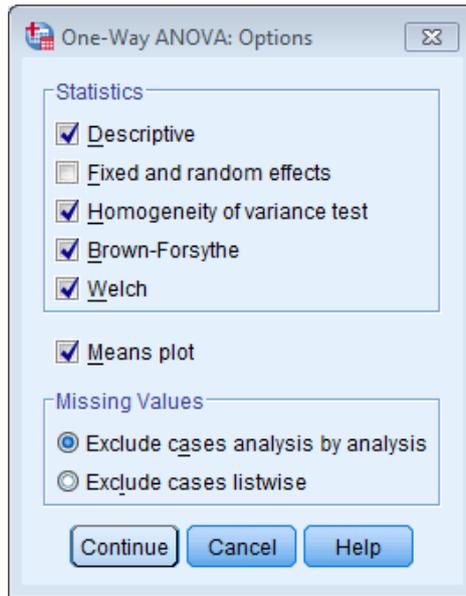


Figure 5: Options for One-Way ANOVA

## Interpreting the Output

Figure 6 shows an error bar chart of the Viagra data (I have edited it). The line indicates a linear trend in that as the dose of Viagra increases so does the mean level of libido.

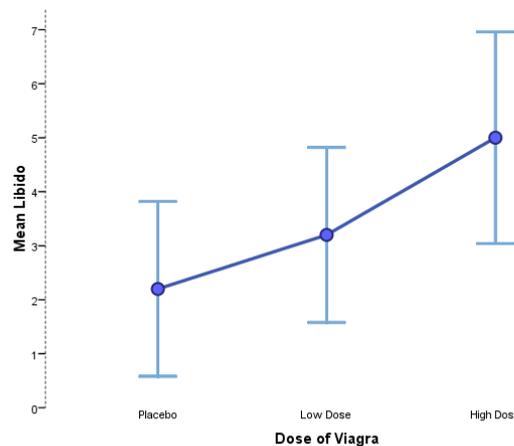
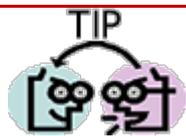


Figure 6: Error bar chart of the Viagra data



- If Levene's test is significant (i.e. the value of sig. is *less than 0.05*) then we can say that the variances are significantly different and the assumption of homogeneity of variance has been violated (see last week and your 'bias' handout).
- Remember that we need to interpret Levene's test within the context of the sample size.

### The main output

Output 1 shows Levene's test (see last week and your handout on bias). For these data the variances are very similar (hence the high probability value), in fact, if you look at some descriptive statistics you'll see that the variances of the placebo and low dose groups are identical.



## Test of Homogeneity of Variances

Libido			
Levene Statistic	df1	df2	Sig.
.092	2	12	.913

### Output 1

Output 2 shows the main ANOVA summary table. The table is divided into between group effects (effects due to the experiment) and within group effects (this is the unsystematic variation in the data). The between group effect is further broken down into a linear and quadratic component and these components are the trend analyses described earlier. The between-group effect labelled *combined* is the overall experimental effect. This overall effect is then broken down because we asked SPSS to conduct trend analyses of these data (we will return to these trends in due course). Had we not specified this, then these two rows of the summary table would not be produced. The row labelled *within group* gives details of the unsystematic variation within the data (the variation due to natural individual differences in libido). The test of whether the group means are the same is represented by the *F*-ratio for the combined between-group effect. The value of this ratio is 5.12. The final column labelled *sig.* Indicates how likely it is that an *F*-ratio of at least that size would have occurred if there were actually no difference between means. In this case, there is a probability of 0.025 (that's only a 2.5% chance!). We can say that there was a significant effect of Viagra because this value is less than the commonly-used threshold of .05.

## ANOVA

Libido			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		20.133	2	10.067	5.119	.025
	Linear Term	Contrast	19.600	1	19.600	9.966	.008
		Deviation	.533	1	.533	.271	.612
	Quadratic Term	Contrast	.533	1	.533	.271	.612
Within Groups			23.600	12	1.967		
Total			43.733	14			

### Output 2

The trend analysis breaks down the experimental effect into that which can be explained by a linear relationship and that which can be explained through a quadratic relationship. First let's look at the linear component. This comparison tests whether the means increase across groups in a linear way. Again the sum of squares and mean squares are given, but the most important things to note are the value of the *F*-ratio and the corresponding significance value. For the linear trend the *F*-ratio is 9.97 and this value is significant at a .008 level of significance. Therefore we can say that as the dose of Viagra increased from nothing to a low dose to a high dose, libido increased proportionately. Moving onto the quadratic trend, this comparison is testing whether the pattern of means is curvilinear (i.e. is represented by a curve with one bend in). The error bar graph of the data strongly suggests that the means cannot be represented by a curve and the results for the quadratic trend bear this out. The *F*-ratio for the quadratic trend is non-significant.

## Output for Planned Comparisons

We told SPSS to conduct two planned comparisons: one to test whether the control group was different to the two groups who received Viagra, and one to see whether the two doses of Viagra made a difference to libido. Output 3 shows the results of the planned comparisons that we requested for the Viagra data. The first table displays the contrast coefficients and it is well worth looking at this table to double check that the contrasts are comparing what they are supposed to. As a quick rule of thumb, remember that when we do planned comparisons we arrange the weights such that we compare any group with a positive weight against any group with a negative weight. Therefore, the table of weights shows that contrast 1 compares the placebo group against the two experimental groups, and contrast 2 compares the low dose group with the high dose group.

The second table gives the statistics for each contrast. The first thing to notice is that statistics are produced for situations in which the group variances are equal, and when they are unequal. If Levene's test was significant then you should read the part of the table labelled *equal variances not assumed*. However, for these data Levene's test was not significant and we can therefore use the part of the table labelled *equal variances assumed*. The table tells us the value of the contrast itself, the associated *t*-test and the two-tailed significance value. Hence, for contrast 1, we can say that taking Viagra significantly increased libido compared to the control group ( $p = .029$ ), but contrast 2 tells us that a high



dose of Viagra did not significantly affect libido compared to a low dose ( $p = .065$ ). Of course contrast 2 is almost significant, which demonstrates my point earlier in this module that significance testing can lead you to very all or nothing thinking. It might be useful to compute some effect sizes.

**Contrast Coefficients**

Contrast	Dose of Viagra		
	Placebo	Low Dose	High Dose
1	-2	1	1
2	0	-1	1

**Contrast Tests**

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Libido	Assume equal variances	1	3.80	1.536	2.474	12	.029
		2	1.80	.887	2.029	12	.065
	Does not assume equal variances	1	3.80	1.483	2.562	8.740	.031
		2	1.80	.917	1.964	7.720	.086

### Output 3



**SELF-TEST:** Compute Cohen's  $d$  for the effect of each group compared to the other. In all cases use the pooled estimate for the SD.<sup>i</sup>

If you calculate  $d$  for the effect of high vs. low doses of Viagra it is  $d = 1.24$ , which is very substantial. Libido in the high dose group is 1.24 standard deviations higher than in the low dose group; this is large effect despite that fact that  $p$  is greater than .05. This fact illustrates how significance testing can cloud what is really going on by drawing your focus onto one aspect of the analysis (i.e. whether  $p$  is greater or smaller than .05). In this case there is clearly something going on between the high and low dose groups that is not reflected in the significance value because of the very small sample (remember that  $p$  is affected by the sample size).

### Output for Post Hoc Tests

If we had no specific hypotheses about the effect that Viagra might have on libido then we could carry out post hoc tests to compare all groups of subjects with each other. In fact, we asked SPSS to do this (see earlier) and the results of this analysis are shown in SPSS Output 4. This table shows the results of Tukey's test (known as Tukey's HSD), the Games-Howell procedure, and Dunnett's test; which were all specified earlier on. If we look at Tukey's test first (because we have no reason to doubt that the population variances are unequal) it is clear from the table that each group of subjects is compared with all of the remaining groups. For each pair of groups the difference between group means is displayed, the standard error of that difference, the significance level of that difference and a 95% confidence interval. First of all, the placebo group is compared to the low dose group and reveals a nonsignificant difference (*Sig.* is greater than .05), but when compared to the high dose group there is a significant difference (*Sig.* is less than .05).

The low dose group is compared to both the placebo group and the high dose group. The first thing to note is that the contrast involving the low dose and placebo group is identical to the one described previously. The only new information is the comparison of the two experimental conditions. The group means differ by 1.8 which is not significant (because  $p = .147$ , which is bigger than .05). The rest of the table describes the Games-Howell tests and a quick inspection reveals the same pattern of results: the only groups that differed significantly were the high dose and placebo groups. These results give us confidence in our conclusions because even if the populations variances are not equal (which seems unlikely given that the sample variances are very similar), then the profile of results still holds true. Finally, Dunnett's test is described and you'll hopefully remember that we asked the computer to compare both experimental groups against the control using a 1-tailed hypothesis that the mean of the control group would be smaller than both experimental groups. Levels of libido in the low dose group are not significantly different (but remember the effect size from earlier) equivalent to the placebo group. However, the high dose group has a significantly higher libido than the placebo group.



## Multiple Comparisons

Dependent Variable: Libido

	(I) Dose of Viagra	(J) Dose of Viagra	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Placebo	Low Dose	-1.000	.887	.516	-3.37	1.37
		High Dose	-2.800*	.887	.021	-5.17	-.43
	Low Dose	Placebo	1.000	.887	.516	-1.37	3.37
		High Dose	-1.800	.887	.147	-4.17	.57
	High Dose	Placebo	2.800*	.887	.021	.43	5.17
		Low Dose	1.800	.887	.147	-.57	4.17
Games-Howell	Placebo	Low Dose	-1.000	.825	.479	-3.36	1.36
		High Dose	-2.800*	.917	.039	-5.44	-.16
	Low Dose	Placebo	1.000	.825	.479	-1.36	3.36
		High Dose	-1.800	.917	.185	-4.44	.84
	High Dose	Placebo	2.800*	.917	.039	.16	5.44
		Low Dose	1.800	.917	.185	-.84	4.44
Dunnnett t (> control) <sup>b</sup>	Low Dose	Placebo	1.000	.887	.227	-.87	
	High Dose	Placebo	2.800	.887	.008	.93	

\*. The mean difference is significant at the 0.05 level.

b. Dunnnett t-tests treat one group as a control, and compare all other groups against it.

Output 4

## Reporting the Results

We saw in the last handout how to report the main ANOVA:

- ✓ There was a significant effect of Viagra on levels of libido,  $F(2, 12) = 5.12, p = .025$ .

Notice that the value of the  $F$ -ratio is preceded by the values of the degrees of freedom for that effect. The linear contrast can be reported in much the same way:

- ✓ There was a significant linear trend,  $F(1, 12) = 9.97, p = .008$ , indicating that as the dose of Viagra increased, libido increased proportionately.

We can also report our effect sizes:

- ✓ Compared to the placebo group, libido in the in the low dose group was about three quarters of a standard deviation higher,  $d = 0.77$ , and nearly two standard deviations higher in the high dose group,  $d = 1.93$ . The high dose group had libido levels more than a standard deviation higher than the low dose group,  $d = 1.24$ . These are all fairly large effects.

Finally, the planned contrasts:

- ✓ Planned contrasts revealed that having any dose of Viagra significantly increased libido compared to having a placebo,  $t(12) = 2.47, p = .029$ , but having a high dose did not significantly increase libido compared to having a low dose,  $t(12) = 2.03, p = .065$ . This later effect is likely to reflect the small sample size because these group means gave rise to a large effect,  $d = 1.24$ .

## Guided Task



Use the data from the one-way ANOVA handout (rugby injuries) that you should have saved in a file called **rugby.sav** to conduct planned comparisons testing the hypotheses that:

- Tonga cause more injuries than all of the other teams.
- Japan cause less injuries than Wales and New Zealand
- Wales and New Zealand are no different in terms of the injuries inflicted.

Did Tonga cause more injuries than all of the other teams (Report some relevant statistics in your answer in APA format)? What are the two values that this contrast is testing the difference between?



**Your Answer:**

Did Japan cause less injuries than Wales and New Zealand (Report some relevant statistics in your answer in APA format)? What are the two values that this contrast is testing the difference between?

**Your Answer:**

Do Wales and New Zealand inflict about the same numbers of injuries (quote relevant statistics in APA format)? What are the two values that this contrast is testing the difference between?

**Your Answer:**



## Unguided Task 1

Use the data from the superhero costume example (that you saved in the file **Superhero.sav**) from the one-way ANOVA handout and test this hypothesis using planned contrasts:

- Costumes of flying superheroes (i.e. Superman and Spiderman) will lead to more severe injuries than non-flying ones (the Hulk and Ninja Turtles)

## Unguided Task 2

Use the data from the soya example (that you saved in the file **sperm.sav**) from the one-way ANOVA handout and test for a linear trend. Also, do some post hoc tests. Which groups differ from which?

## Unguided Task 3

Use the data from the mobile phone example (that you saved in the file **tumour.sav**) from the one-way ANOVA handout and test for a linear trend. Also, design some planned comparisons and test them.

## Multiple Choice Test



Complete the multiple choice questions for **Chapter 10** on the companion website to Field (2013): <https://studysites.uk.sagepub.com/field4e/study/mcqs.htm>. If you get any wrong, re-read this handout and the one from last week (or Field, 2013, Chapter 11) and do them again until you get them all correct.

## References

Field, A. P. (2013). *Discovering statistics using IBM SPSS Statistics: And sex and drugs and rock 'n' roll* (4th ed.). London: Sage.

## Terms of Use

This handout contains material from:

Field, A. P. (2013). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll* (4th Edition). London: Sage.

This material is copyright Andy Field (2000-2016).

This document is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/), basically you can use it for teaching and non-profit activities but not meddle with it without permission from the author.

---

<sup>i</sup> For Low vs. Placebo you should get  $d = 0.77$ , which is fairly large; for High vs. Placebo,  $d = 1.93$ , which is very high; and for Low vs. High,  $d = 1.24$ , which is very substantial.