

Defining Contrasts with Syntax

Why Do We Need To Use Syntax?

In Chapters 10, 11 and 12 of *Discovering Statistics Using SPSS* we used SPSS's built in contrast functions to compare various groups after conducting ANOVA. These special contrasts (described in Chapter 8, Table 8.6) cover many situations, but in more complex designs there will be times when you want to do contrasts that simply can't be done using SPSS's built in contrasts. Unlike One-Way ANOVA, there is no way in factorial designs to define contrast codes through the Windows dialog boxes. However, SPSS can do these contrasts if you define them using syntax.

An Example

Imagine a clinical psychologist wanted to see the effects of a new antidepressant drug called Cheerup. He took 50 people suffering from clinical depression and randomly assigned them to one of five groups. The first group were a waiting list control group (that is, they were people assigned to the waiting list who were not treated during the study), the second took a placebo tablet (i.e. they were told they were being given an antidepressant drug but actually the pills contained sugar and no active agents), the third group took a well-established SSRI antidepressant called Seroxat (Paxil to American readers), the fourth group were given a well established SNRI antidepressant called Effexor¹, the final group were given the new drug, Cheerup. Levels of depression were measured before and after 2 months on the various treatments, and ranged from 0 = as happy as a spring lamb to 20 = pass me the noose. The data are in the file **Depression.sav**.

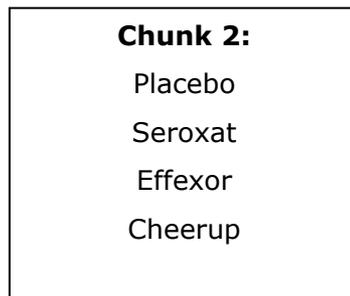
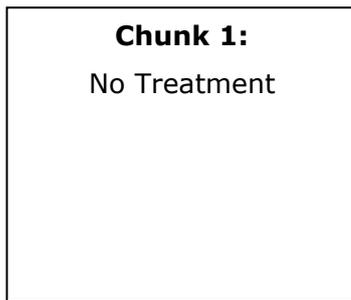
The design of this study is a two-way Mixed design. There are two independent variables: treatment (no treatment, placebo, Seroxat, Effexor or Cheerup), and time (before or after treatment). Treatment is measured with different participants (and so is between-group) and time is, obviously, measured using the same participants (and so is repeated measures). Hence, the ANOVA we want to use is a 5×2 two-way ANOVA.

Now, we want to do some contrasts. Imagine we have the following hypotheses:

1. Any treatment will be better than no treatment
2. Drug Treatments will be better than the placebo
3. Our new drug, Cheerup will be better than old-style antidepressants
4. The old-style antidepressants will not differ in their effectiveness.

We have to code these various hypotheses as we did in chapter 8. The first contrast involves comparing the no treatment condition with all other groups. Therefore, the first step is to chunk these variables, and then assign a positive weight to one chunk and a negative weight to the other chunk.

¹ SSRIs are Selective Serotonin Reuptake Inhibitors and so work selectively inhibit the reuptake of the neurotransmitter serotonin in the brain, whereas SNRIs are the newer Serotonin-Norepinephrine Reuptake Inhibitors, which act not only on serotonin but on another neurotransmitter (from the same family), Norepinephrine. If you're interested in depression then can I shamelessly plug my book on clinical psychology (Field, A. P., (2003). *Clinical Psychology*. Crucial: Exeter. See www.learningmatters.co.uk)!

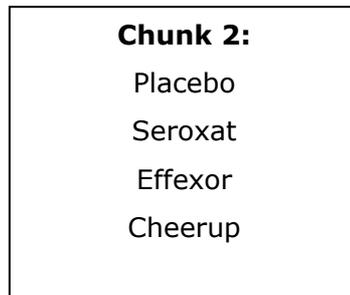
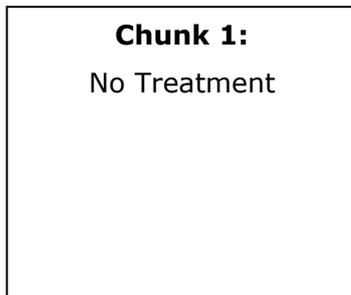


-

+

Sign of Weight

Having done that, we need to assign a numeric value to the groups in each chunk. As I mentioned in chapter 8, the easiest way to do this is just to assign a value equal to the number of groups in the *opposite* chunk. Therefore, the value for any group in chunk 1 will be the same as the number of groups in chunk 2 (in this case 4). Likewise, the value for any groups in chunk 2 will be the same as the number of groups in chunk 1 (in this case 1). So, we get the following codes:



-

+

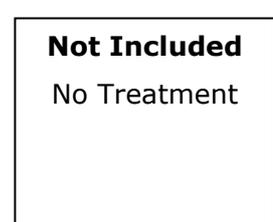
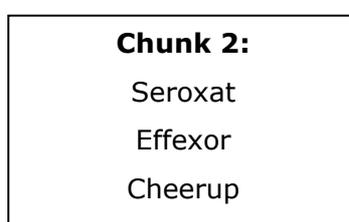
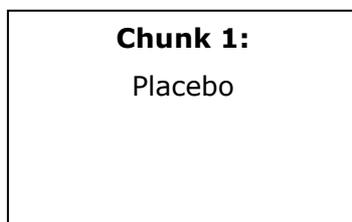
Sign of Weight

-4

+1

Value of Weight

The second contrast requires us to compare the Placebo group to all of the drug groups. Again, we chunk our groups accordingly, assign one chunk a negative sign and the other a positive and then assign a weight on the basis of the number of groups in the opposite chunk. We must also remember to give the no treatment group a weight of 0 because they're not involved in the contrast.



-

+

Sign of Weight

-3

+1

Value of Weight

0

The third contrast requires us to compare the new drug (Cheerup) with the old drugs (Seroxat and Effexor). Again, we chunk our groups accordingly, assign one chunk a negative sign and

the other a positive and then assign a weight on the basis of the number of groups in the opposite chunk. We must also remember to give the no treatment and placebo groups a weight of 0 because they're not involved in the contrast.

| | | |
|----------------------------|---------------------------------------|------------------------------------------------|
| Chunk 1: Cheerup | Chunk 2: Seroxat Effexor | Not Included No Treatment Placebo |
| - | + | Sign of Weight |
| -2 | +1 | Value of Weight |
| | | 0 |

The final contrast requires us to compare the two old drugs. Again, we chunk our groups accordingly, assign one chunk a negative sign and the other a positive and then assign a weight on the basis of the number of groups in the opposite chunk. We must also give the no treatment, placebo, and Cheerup groups a weight of 0 because they're not involved in the contrast.

| | | |
|----------------------------|----------------------------|-----------------------------------------------------------|
| Chunk 1: Effexor | Chunk 2: Seroxat | Not Included No Treatment Placebo Cheerup |
| - | + | Sign of Weight |
| -1 | +1 | Value of Weight |
| | | 0 |

We can summarise these codes in the following table:

| | No Treatment | Placebo | Seroxat | Effexor | Cheerup |
|------------|--------------|---------|---------|---------|---------|
| Contrast 1 | -4 | 1 | 1 | 1 | 1 |
| Contrast 2 | 0 | -3 | 1 | 1 | 1 |
| Contrast 3 | 0 | 0 | 1 | 1 | -2 |
| Contrast 4 | 0 | 0 | 1 | -1 | 0 |

These are the codes that we need to enter into SPSS to do the contrasts that we'd like to do.

Entering the Contrasts Using Syntax

To enter these contrasts using syntax we have to first open a syntax window (see Chapter 2 of Discovering Statistics). Having done that we have to type the following commands:

```
MANOVA
before after BY treat(0 4)
```

This initialises the ANOVA command in SPSS. The second line specifies the variables in the data editor. The first two words *before* and *after* are the repeated measures variables (and these words are the words used in the data editor). Anything after *BY* is a between group measures and so needs to be followed by brackets

within which the minimum and maximum values of the coding variable are specified. I called the between-group variable `treat`, and I coded the groups as 0 = no treatment, 1 = placebo, 2 = Seroxat, 3 = Effexor, 4 = Cheerup. Therefore, the minimum and maximum codes were 0 and 4. So these two lines tell SPSS to start the ANOVA procedure, that there are two repeated measures variables called before and after and that there is a between group variable called treat that has a minimum code of 0 and a maximum of 4.

```
/WSFACTORS time (2)
```

The `/WSFACTORS` command allows us to specify any repeated measures variables. SPSS already knows that there are two variables called before and after, but it doesn't know how to treat these variables. This command tells SPSS to create a repeated measures variables called `time` that has two levels (the number in brackets). SPSS then looks to the variables specified before and assigns the first one (`before` in this case) to be the first level of time, and then assigns the second one (in this case `after`) to be the second level of `time`.

```
/CONTRAST (time)=special(1 1, 1 -1)
```

This is used to specify the contrasts for the first variable. The `/CONTRAST` is used to specify any contrast. It's always followed by the name of the variable that you want to do a contrast on in brackets. We have two variables (`time` and `treat`) and in this first contrast we want to specify a contrast for `time`. Time only has two levels and so all we want to do is to tell SPSS to compare these two levels (which actually it will do by default but I want you to get some practice in!). What we write after the equals sign defines the contrast, so we could write the name of one of the standard contrasts such as `Helmert`, but because we want to specify our own contrast we use the word `special`. Special should always be followed by brackets, and inside those brackets are your contrast codes. Codes for different contrasts are separated using a comma, and within a contrast, codes for different groups are separated using a space. The first contrast should always be one that defines a baseline for all other contrasts and that is one that codes all groups with a 1. Therefore, because we have two levels of time, we just write `1 1`, which tells SPSS that the first contrast should be one in which both before and after are given a code of 1. The comma tells SPSS that a new contrast follows and this second contrast has been defined as `1 -1` and this tells SPSS that in this second contrast we want to give `before` a code of 1, and `after` a code of -1. Note that the codes you write in the brackets are assigned to variables in the order that those variables are entered into the SPSS Syntax, so because we originally wrote `before after BY treat(0 4)` SPSS assigns the 1 to `before` and -1 to `after`, if we'd originally wrote `after before BY treat(0 4)` then SPSS would have assigned them the opposite way round: the 1 to `after` and -1 to `before`.

```
/CONTRAST (treat)=special (1 1 1 1 1, -4 1 1 1 1, 0 -3 1 1 1, 0 0 1 1 -2, 0 0 1 -1 0)
```

This is used to specify the contrasts for the second variable. This time the `/CONTRAST` command is followed by the name of the second variable (`treat`) variable. Treat has five levels and we've already worked out 4 different contrasts that we want to do. Again we use the word `special` after the equals sign and specify our coding values within the brackets. As before, codes for different contrasts are separated using a comma, and within a contrast, codes for different groups are separated using a space. Also, as before, the first contrast should always be one that defines a baseline for all other contrasts and that is one that codes all groups with a 1. Therefore, because we have five levels of time, we just write `1 1 1 1 1`, which tells SPSS that the first contrast should be one in which all five groups are given a code of 1. The comma tells SPSS that a new contrast follows and this second contrast has been defined as `-4 1 1 1 1`, and this tells SPSS that in this second contrast we want to give the first group a code of -4, and all subsequent groups codes of 1. How does SPSS decide what the first group is? It uses the coding variable in the data editor and orders the groups in the same order as the coding variable. Therefore, because I coded the groups as 0 = no treatment, 1 = placebo, 2 = Seroxat, 3 = Effexor, 4 = Cheerup, this first contrast gives the no treatment group a code of -4, and all subsequent groups codes of 1. The comma again tells SPSS that having done this, there is another contrast to follow and this contrast has been defined as `0 -3 1 1 1`, and this tells SPSS that in this contrast we want to give the first group a code of 0, the second group a code of -3 and all subsequent groups codes of 1. Again, because I coded the groups as 0 = no treatment, 1 = placebo, 2 = Seroxat, 3 = Effexor, 4 = Cheerup, this contrast gives the no treatment group a code of 0, the placebo group a code of -3 and all subsequent groups codes of 1. The comma again tells SPSS that having done this, there is another contrast to follow and this contrast has been defined as `0 0 1 1 -2`, and this tells SPSS that in this contrast we want to give the first two groups a code of 0, the third and fourth group a code of 1 and the final group a code of -2. Again, because I coded the groups as 0 = no treatment, 1 = placebo, 2 = Seroxat, 3 = Effexor, 4 =

= Cheerup, this contrast gives the no treatment and placebo groups a code of 0, the Effexor and Seroxat groups a code of -3 and the Cheerup group a code of 1. The comma again tells SPSS that there is yet another contrast to follow and this contrast has been defined as 0 0 1 -1 0, and this tells SPSS that in this contrast we want to give the first, second and last groups a code of 0, the third group a code of 1 and the fourth group a code of -1. This contrast gives the no treatment, placebo and Cheerup groups a code of 0, the Seroxat group a code of 1 and the Effexor group a code of -1. As such, this one line of text has defined the four contrasts that we want to do.

```
/CINTERVAL JOINT(.95) MULTIVARIATE(BONFER)
```

This line defines the type of confidence intervals that you want to do for your contrasts. I recommend the Bonferroni option, but if you delve into the SPSS syntax guide you can find others.

```
/METHOD UNIQUE  
/ERROR WITHIN+RESIDUAL  
/PRINT TRANSFORM HOMOGENEITY(BARTLETT COCHRAN BOXM)  
SIGNIF( UNIV MULT AVERF HF GG )  
PARAM( ESTIM EFSIZE).
```

These lines of syntax specify various things (that may or may not be useful) such as a transformation matrix (TRANSFORM), which isn't at all necessary here but is useful if you've used SPSS's built-in contrasts, Homogeneity tests (HOMOGENEITY(BARTLETT COCHRAN BOXM)), the main ANOVA and Huynh-Feldt and Greenhouse-Geisser corrections which we don't actually need in this example (SIGNIF(UNIV MULT AVERF HF GG)), and parameter estimates and effect size estimates for the contrasts we've specified (PARAM(ESTIM EFSIZE)) line defines the type of confidence intervals that you want to do for your contrasts. So, the whole syntax will look like this:

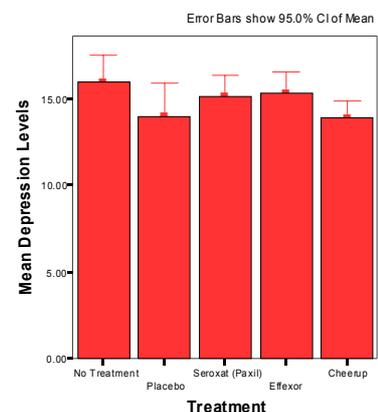
```
MANOVA  
before after BY treat(0 4)  
/WSFACTORS time (2)  
/CONTRAST (time)=special(1 1, 1 -1)  
/CONTRAST (treat)=special (1 1 1 1 1, -4 1 1 1 1, 0 -3 1 1 1, 0 0 1 1 -2, 0 0 1 -1 0)  
/CINTERVAL JOINT(.95) MULTIVARIATE(BONFER)  
/METHOD UNIQUE  
/ERROR WITHIN+RESIDUAL  
/PRINT TRANSFORM HOMOGENEITY(BARTLETT COCHRAN BOXM)  
SIGNIF( UNIV MULT AVERF HF GG )  
PARAM( ESTIM EFSIZE).
```

It's very important to remember the full stop at the end! This syntax is in the file **DepressionSyntax.sps** as well in case your typing goes wrong!

Output From The Contrasts

The output you get is in the form of text (no nice pretty tables) and to interpret it you have to remember the contrasts you specified! I'll run you through the main highlights of this example:

The first bit of the output will show the homogeneity tests (which should all be non-significant but beware of Box's test because it tends to be inaccurate). The first important part is the main effect of the variable treat. First there's an ANOVA summary table like those you've come across before (if you've read Chapters 8-11). This tells us that there's no significant main effect of the type of treatment, $F(4, 45) =$



2.01, $p > .05$. This tells us that if you ignore the time at which depression was measured then the levels of depression were about the same across the treatment groups. Of course, levels of depression should be the same before treatment, and so this isn't a surprising result (because it averages across scores before and after treatment). The graph shows that, in fact levels of depression are relatively similar across groups.

***** Analysis of Variance -- design 1 *****

Tests of Between-Subjects Effects.

| Tests of Significance for T1 using UNIQUE sums of squares | | | | | |
|-----------------------------------------------------------|--------|----|-------|------|----------|
| Source of Variation | SS | DF | MS | F | Sig of F |
| WITHIN+RESIDUAL | 359.95 | 45 | 8.00 | | |
| TREAT | 64.30 | 4 | 16.08 | 2.01 | .109 |

Estimates for T1

--- Joint univariate .9500 BONFERRONI confidence intervals

TREAT

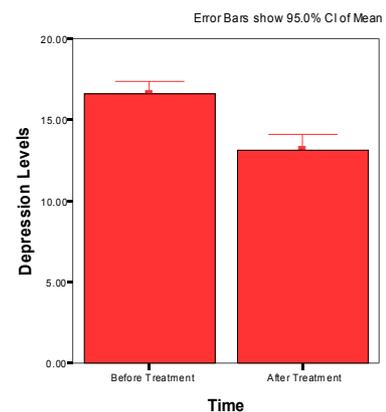
| Parameter | Coeff. | Std. Err. | t-Value | Sig. t | Lower -95% CL | Upper |
|-----------|------------|-----------|----------|--------|---------------|----------|
| 2 | -7.7781746 | 3.99972 | -1.94468 | .05808 | -18.18578 | 2.62944 |
| 3 | 3.53553391 | 3.09817 | 1.14117 | .25984 | -4.52617 | 11.59723 |
| 4 | 3.74766594 | 2.19074 | 1.71069 | .09403 | -1.95282 | 9.44815 |
| 5 | -.21213203 | 1.26482 | -.16772 | .86756 | -3.50331 | 3.07904 |

Parameter ETA Sq.

| | |
|---|--------|
| 2 | .07752 |
| 3 | .02813 |
| 4 | .06106 |
| 5 | .00062 |

This main effect is followed by some contrasts, but we don't need to look at these because the main effect was non-significant. However, just to tell you what they are, parameter 2 is our first contrast (no treatment vs. the rest) and as you can see this is almost significant (p is just above 0.05), parameter 3 is our second contrast (placebo vs. the rest) and this is non-significant. Parameter 4 is our third contrast (Cheerup vs. Effexor and Seroxat), and again this is almost significant. Parameter 5 is our last contrast (Seroxat vs. Effexor) and this is very non-significant. However, these contrasts all ignore the effect of time and so aren't really what we're interested in.

The next part that we're interested in is the within-subjects effects, and this involves the main effect of Time and the interaction of time and treatment. First there's an ANOVA summary table as before. This tells us that there's a significant main effect of the time, $F(1, 45) = 43.02$, $p < .001$. This tells us that if you ignore the type of treatment, there was a significant difference between depression levels before and after treatment. A quick look at the means reveals that depression levels were significantly lower after treatment. Below the ANOVA table is a parameter estimate for the effect of time. As there are only 2 levels of time, this represents the difference in depression levels before and after treatment. No other contrasts are possible.



***** Analysis of Variance -- design 1*****

Tests involving 'TIME' Within-Subject Effect.

| Source of Variation | SS | DF | MS | F | Sig of F |
|---------------------|--------|----|--------|-------|----------|
| WITHIN+RESIDUAL | 320.35 | 45 | 7.12 | | |
| TIME | 306.25 | 1 | 306.25 | 43.02 | .000 |
| TREAT BY TIME | 125.90 | 4 | 31.47 | 4.42 | .004 |

Estimates for T2
 --- Joint univariate .9500 BONFERRONI confidence intervals

TIME

| Parameter | Coeff. | Std. Err. | t-Value | Sig. | t Lower -95% | CL- Upper |
|-----------|------------|-----------|---------|--------|--------------|-----------|
| 1 | 2.47487373 | .37733 | 6.55891 | .00000 | 1.71489 | 3.23485 |

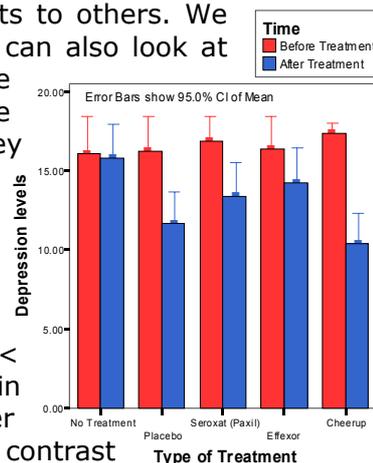
| Parameter | ETA Sq. |
|-----------|---------|
| 1 | .48875 |

TREAT BY TIME

| Parameter | Coeff. | Std. Err. | t-Value | Sig. | t Lower -95% | CL- Upper |
|-----------|------------|-----------|----------|--------|--------------|-----------|
| 2 | 11.3137085 | 3.77330 | 2.99836 | .00441 | 1.49527 | 21.13214 |
| 3 | -.56568542 | 2.92278 | -.19354 | .84740 | -8.17101 | 7.03964 |
| 4 | -5.8689863 | 2.06672 | -2.83976 | .00675 | -11.24676 | -.49121 |
| 5 | .919238816 | 1.19322 | .77038 | .44510 | -2.18562 | 4.02410 |

| Parameter | ETA Sq. |
|-----------|---------|
| 2 | .16651 |
| 3 | .00083 |
| 4 | .15197 |
| 5 | .01302 |

The interaction term is also significant, $F(4, 45) = 4.42, p < .01$. This indicates that the change in depression over time is different in some treatments to others. We can make sense of this through an interaction graph, but we can also look at our contrasts. The key contrasts for this whole analysis are the parameter estimates for the interaction term (the bit in the output underneath the heading TREAT BY TIME) because they take into account the effect of time *and* treatment.



- ✓ Parameter 2 is our first contrast (no treatment vs. the rest) and as you can see this is significant (p is below 0.05). This tells us that the change in depression levels in the no treatment group was significantly different to the average change in all other groups, $t = 2.998, p < .01$. As you can see in the graph, there is no change in depression in the no treatment group, but in all other groups there is a fall in depression. Therefore, this contrast reflects the fact that there is no change in the no treatment group, but there is a decrease in depression levels in all other groups.
- ✓ Parameter 3 is our second contrast (placebo vs. Seroxat, Effexor and Cheerup) and this is very non-significant, $t = -.194, p = .85$. This shows that the decrease in depression levels seen in the placebo group is comparable to the average decrease in depression levels seen in the Seroxat, Effexor and Cheerup conditions. In other words, the combined effect of the drugs on depression is no better than a placebo.

- ✓ Parameter 4 is our third contrast (Cheerup vs. Effexor and Seroxat), and again this is highly significant, $t = -2.84$, $p < 0.01$. This shows that the decrease in depression levels seen in the Cheerup group is significantly bigger than the decrease seen in the Effexor and Seroxat groups combined. Put another way Cheerup has a significantly bigger effect than other established antidepressants.
- ✓ Parameter 5 is our last contrast (Seroxat vs. Effexor) and this is very non-significant, $t = .77$, $p = 0.45$. This tells us that the decrease in depression levels seen in the Seroxat group is comparable to the decrease in depression levels seen in the, Effexor group. Put another way, Effexor and Seroxat seem to have similar effects on depression.

I hope to have shown in this example how to specify contrasts using syntax and how looking at these contrasts (especially for an interaction term) can be a very useful way to break down an interaction effect.

This document is distributed on the CD-ROM of:

Field, A. P. (2004). *Discovering Statistics Using SPSS (2nd Edition)*. London: Sage.

It should be referenced accordingly, and is copyright protected with the rest of the book.