

Chapter 20: Multilevel linear models

Oliver Twisted

Please, Sir, can I have some more ... ICC?

The following article appears in:



Field, A. P. (2005). Intraclass correlation. In B. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 2, pp. 948–954). Hoboken, NJ: Wiley.

It appears in adapted form below.

Commonly used correlations such as the *Pearson product moment correlation* measure the bivariate relation between variables of different measurement classes. These are known as *interclass* correlations. By ‘different measurement classes’ we really just mean variables measuring different things. For example, we might look at the relation between attractiveness and career success. Clearly one of these variables represents a class of measures of how good looking a person is, whereas the other represents the class of measurements of something quite different: how much someone achieves in their career. However, there are often cases in which it is interesting to look at relations between variables *within* classes of measurement. In its simplest form, we might compare only two variables. For example, we might be interested in whether anxiety runs in families, and we could look at this by measuring anxiety within pairs of twins (Eley & Stevenson, 1999). In this case the objects being measured are twins, and both twins are measured on some index of anxiety. As such, there is a pair of variables both measuring anxiety, therefore, from the same class. In such cases, an *intraclass* correlation (ICC) is used and is commonly extended beyond just two variables to look at the consistency between judges. For example, in gymnastics, ice-skating, diving and other Olympic sports, contestants’ performance is often assessed by a panel of judges. There might be 10 judges, all of whom rate performance out of 10; therefore, the resulting measures are from the same class (they measure the same thing). The objects being rated are the competitors. This again is a perfect scenario of an intraclass correlation.

Models of intraclass correlations

There are a variety of different intraclass correlations (McGraw & Wong, 1996; Shrout & Fleiss, 1979), and the first step in calculating one is to determine a model for your sample data. All of the various forms of the intraclass correlation are based on estimates of mean variability from a one-way repeated-measures *analysis of variance*.

All situations in which an intraclass correlation is desirable will involve multiple measures on different entities (be they twins, Olympic competitors, pictures, sea slugs, etc.). The objects measured constitute a random factor in the design (they are assumed to be random exemplars of the population of objects). The measures taken can be included as factors in the design if they have a meaningful order, or can be excluded if they are unordered, as we shall now see.

One-way random effects model

In the simplest case we might have only two measures (think back to our twin study on anxiety). When the order of these variables is irrelevant (for example, with our twins it is arbitrary whether we treat the data from the first twin as being anxiety measure 1 or anxiety measure 2). In this case, the only systematic source of variation is the random variable representing the different objects. As such, we can use a one-way ANOVA of the form:

$$x_{ij} = \mu + r_i + e_{ij}$$

in which r_i is the effect of object i (known as the row effects), j is the measure being considered, and e_{ij} is an error term (the residual effects). The row and residual effects are random, independent and normally distributed. Because the effect of the measure is ignored, the resulting intraclass correlation is based on the overall effect of the objects being measured (the mean between-object variability, MS_{Rows}) and the mean within-object variability (MS_{W}). Both of these will be formally defined later.

Two-way random effects model

When the order of measures is important then the effect of the measures becomes important. The most common case of this is when measures come from different judges or raters. Hodgins and Makarchuk (2003), for example, show two such uses; in their study they took multiple measures of the same class of behaviour (gambling) but also took measures from different sources. They measured gambling both in terms of days spent gambling and money spent gambling. Clearly these measures generate different data, so it is important to which measure a datum belongs (it is not arbitrary to which measure a datum is assigned). This is one scenario in which a two-way model is used. However, they also took measures of gambling both from the gambler and a collateral (e.g., spouse). Again, it is important that we attribute data to the correct source. So, this is a second illustration of where a two-way model is useful. In such situations the intraclass correlation can be used to check the consistency or agreement between measures or raters. In this situation a two-way model can be used as follows:

$$x_{ij} = \mu + r_i + c_j + rc_{ij} + e_{ij}$$

in which c_j is the effect of the measure (i.e., the effect of different raters, or different measures), and rc_{ij} is the interaction between the measures taken and the objects being measured. The effect of the measure (c_j) can be treated as either a fixed effect or a random effect. How it is treated doesn't affect the calculation of the intraclass correlation, but it does affect the interpretation (as we shall see). It is also possible to exclude the interaction term and use the model:

$$x_{ij} = \mu + r_i + c_j + e_{ij}$$

We shall now turn our attention to calculating the sources of variance needed to calculate the intraclass correlation.

Sources of variance: an example

In the chapter in the book on repeated-measures ANOVA, there is an example relating to student concerns about the consistency of marking between lecturers. It is common that lecturers obtain reputations for being 'hard' or 'light' markers, which can lead students to believe that their marks are not based solely on the intrinsic merit of the work, but can be influenced by who marked the work. To test this we could calculate an intraclass correlation. First, we could submit the same eight essays to four different lecturers and record the mark they gave each essay. Table 1 shows the data, and you should note that it looks the same as a one-way repeated-measures ANOVA in which the four lecturers represent four levels of an 'independent variable' and the outcome or dependent variable is the mark given (in fact I use these data as an example of a one-way repeated-measures ANOVA).

Table 1

Essay	Dr Field	Dr Smith	Dr Scrote	Dr Death	Mean	S^2	$S^2(k-1)$
1	62	58	63	64	61.75	6.92	20.75
2	63	60	68	65	64.00	11.33	34.00
3	65	61	72	65	65.75	20.92	62.75
4	68	64	58	61	62.75	18.25	54.75
5	69	65	54	59	61.75	43.58	130.75
6	71	67	65	50	63.25	84.25	252.75
7	78	66	67	50	65.25	132.92	398.75
8	75	73	75	45	67.00	216.00	648.00
Mean:	68.88	64.25	65.25	57.38	63.94	Total:	1602.50

There are three different sources of variance that are needed to calculate an intraclass correlation which we shall now calculate. These sources of variance are the same as those

calculated in one-way repeated-measures ANOVA. (If you don't believe me, consult Smart Alex's answers to Chapter 13 to see an identical set of calculations!)

The between-object variance (MS_{Rows})

The first source of variance is the variance between the objects being rated (in this case the between-essay variance). Essays will naturally vary in their quality for all sorts of reasons (the natural ability of the author, the time spent writing the essay, etc.). This variance is calculated by looking at the average mark for each essay and seeing how much it deviates from the average mark for all essays. These deviations are squared because some will be positive and others negative and so would cancel out when summed. The squared errors for each essay are weighted by the number of values that contribute to the mean (in this case the number of different markers, k). So, in general terms we write this as:

$$SS_{\text{Rows}} = \sum_{i=1}^n k_i (\bar{X}_{\text{Row } i} - \bar{X}_{\text{all rows}})^2$$

Or, for our example, we could write it as:

$$SS_{\text{Essays}} = \sum_{i=1}^n k_i (\bar{X}_{\text{Essay } i} - \bar{X}_{\text{all essays}})^2$$

This would give us:

$$\begin{aligned} SS_{\text{Rows}} &= 4(61.75 - 63.94)^2 + 4(64.00 - 63.94)^2 + 4(65.75 - 63.94)^2 + 4(62.75 - 63.94)^2 \\ &\quad + 4(61.75 - 63.94)^2 + 4(63.25 - 63.94)^2 + 4(65.25 - 63.94)^2 + 4(67.00 - 63.94)^2 \\ &= 19.18 + 0.014 + 13.10 + 5.66 + 19.18 + 1.90 + 6.86 + 37.45 \\ &= 103.34 \end{aligned}$$

This sum of squares is based on the total variability and so its size depends on how many objects (essays in this case) have been rated. Therefore, we convert this total to an average known as the mean squared error (MS) by dividing by the number of essays (or in general terms the number of rows) minus 1. This value is known as the *degrees of freedom*.

$$MS_{\text{Rows}} = \frac{SS_{\text{Rows}}}{df_{\text{Rows}}} = \frac{103.34}{n-1} = \frac{103.34}{7} = 14.76$$

The mean squared error for the rows in the table is our estimate of the natural variability between the objects being rated.

The within-judge variability (MS_{W})

The second variability in which we're interested is the variability within measures/judges. To calculate this we look at the deviation of each judge from the average of all judges on a particular essay. We use an equation with the same structure as before, but for each essay separately:

$$SS_{\text{Essay}} = \sum_{k=1}^p (\bar{X}_{\text{Column } k} - \bar{X}_{\text{all columns}})^2$$

For essay 1, for example, this would be:

$$\begin{aligned} SS_{\text{Essay1}} &= (62 - 61.75)^2 + (58 - 61.75)^2 + (63 - 61.75)^2 + (64 - 61.75)^2 \\ &= 20.75 \end{aligned}$$

The degrees of freedom for this calculation are again one less than the number of scores used in the calculation. In other words it is the number of judges, k , minus 1.

We have to calculate this for each of the essays in turn and then add these values up to get the total variability within judges. An alternative way to do this is to use the variance within each essay. The equation mentioned above is equivalent to the variance for each essay multiplied by the number of values on which that variance is based (in this case the number of judges, k) minus 1. As such we get:

$$SS_w = s_{\text{Essay 1}}^2 (k_1 - 1) + s_{\text{Essay 2}}^2 (k_2 - 1) + \dots + s_{\text{Essay } n}^2 (k_n - 1)$$

Table 1 shows the values for each essay in the last column. When we sum these values we get 1602.50. As before, this value is a total and so depends on the number essays (and the number of judges). Therefore, we convert it to an average, by dividing by the degrees of freedom. For each essay we calculated a sum of squares that we saw was based on $k - 1$ degrees of freedom. Therefore, the degrees of freedom for the total within-judge variability are the sum of the degrees of freedom for each essay:

$$df_w = n(k - 1)$$

in which n is the number of essays and k is the number of judges. In this case it will be $8(4 - 1) = 24$.

The resulting mean squared error is, therefore:

$$MS_w = \frac{SS_w}{df_w} = \frac{1602.50}{n(k - 1)} = \frac{1602.50}{24} = 66.77$$

The between-judge variability (MS_{Columns})

The within-judge or within-measure variability is made up of two components. The first is the variability created by *differences* between judges. The second is unexplained variability (error for want of a better word). The variability between judges is again calculated using a variant of the same equation that we've used all along, only this time we're interested in the deviation of each judge's mean from the mean of all judges:

$$SS_{\text{Columns}} = \sum_{k=1}^p n_i (\bar{X}_{\text{Column } i} - \bar{X}_{\text{all columns}})^2$$

or

$$SS_{\text{Judges}} = \sum_{k=1}^p n_i (\bar{X}_{\text{Judge } i} - \bar{X}_{\text{all judges}})^2$$

in which n is the number of things that each judge rated. For these data we'd get:

$$\begin{aligned} SS_{\text{Columns}} &= 8(68.88 - 63.94)^2 + 8(64.25 - 63.94)^2 + 8(65.25 - 63.94)^2 + 8(57.38 - 63.94)^2 \\ &= 554 \end{aligned}$$

The degrees of freedom for this effect are the number of judges, k , minus 1. As before, the sum of squares is converted to a mean squared error by dividing by the degrees of freedom:

$$MS_{\text{Columns}} = \frac{SS_{\text{Columns}}}{df_{\text{Columns}}} = \frac{554}{k-1} = \frac{554}{3} = 184.67$$

The error variability (MS_E)

The final variability is the variability that can't be explained by known factors such as variability between essays or judges/measures. This can be easily calculated using subtraction because we know that the within-judges variability is made up of the between-judges variability and this error:

$$\begin{aligned} SS_W &= SS_{\text{Columns}} + SS_E \\ SS_E &= SS_W - SS_{\text{Columns}} \end{aligned}$$

The same is true of the degrees of freedom:

$$\begin{aligned} df_W &= df_{\text{Columns}} + df_E \\ df_E &= df_W - df_{\text{Columns}} \end{aligned}$$

So, for these data we get:

$$SS_E = SS_W - SS_{\text{Columns}} = 1602.50 - 554 = 1048.50$$

and

$$df_E = df_W - df_{\text{Columns}} = 24 - 3 = 21$$

We get the average error variance in the usual way:

$$MS_E = \frac{SS_E}{df_E} = \frac{1048.50}{21} = 49.93$$

Calculating intraclass correlations

Having computed the necessary variance components, we shall now look at how the intraclass correlation is calculated. Before we do so, however, there are two important decisions to be made.

Single measures or average measures

So far we have talked about situations in which the measures we've used produce single values. However, it is possible that we might have measures that produce an average score. For example, we might get judges to rate paintings in a competition based on style, content, originality, and technical skill. For each judge, their ratings are averaged. The end result is still ratings from a set of judges, but these ratings are an average of many ratings. Intraclass correlations can be computed for such data, but the computation is somewhat different.

Consistency or agreement?

The next decision involves whether you want a measure of overall consistency between measures/judges. The best way to explain this distinction is to return to our lecturers marking essays. It is possible that particular lecturers are harsh in their ratings (or lenient). A consistency definition views these differences as an irrelevant source of variance. As such the between-judge variability described above ($MS_{Columns}$) is ignored in the calculation (see Table 2). In ignoring this source of variance we are getting a measure of whether judges agree about the relative merits of the essays without worrying about whether the judges anchor their marks around the same point. So, if all the judges agree that essay 1 is the best, essay 5 is the worst (or their rank order of essays is roughly the same) then agreement will be high: it doesn't matter that Dr Field's marks are all 10% higher than Dr Death's. This is a consistency definition of agreement.

The alternative is to treat relative differences between judges as an important source of disagreement. That is, the between-judge variability described above ($MS_{Columns}$) is treated as an important source of variation and is included in the calculation (see Table 2). In this scenario disagreements between the relative magnitude of judges' ratings matter (so, the fact that Dr Death's marks differ from Dr Field's will matter even if their rank order of marks is in agreement). This is an absolute agreement definition. By definition the one-way model ignores the effect of the measures and so can have only this kind of interpretation.

Equations for ICCs

Table 2 shows the equations for calculating ICC based on whether a one-way or two-way model is assumed and whether a consistency or absolute agreement definition is preferred. For illustrative purposes, the ICC is calculated in each case for the example used in this entry. This should enable the reader to identify how to calculate the various sources of variance. In this table $MS_{Columns}$ is abbreviated to MS_C and MS_{Rows} is abbreviated to MS_R .

Table 2

ICC for Single Scores			
Model	Interpretation	Equation	ICC for example data
One-way	Absolute agreement	$\frac{MS_R - MS_W}{MS_R + (k-1)MS_W}$	$\frac{14.76 - 66.77}{14.76 + (4-1)66.77} = -0.24$
Two-way	Consistency	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E}$	$\frac{14.76 - 49.93}{14.76 + (4-1)49.93} = -0.21$
	Absolute agreement	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$	$\frac{14.76 - 49.93}{14.76 + (4-1)49.93 + \frac{4}{8}(184.67 - 49.93)} = -0.15$
ICC for Average Scores			
One-way	Absolute agreement	$\frac{MS_R - MS_W}{MS_R}$	$\frac{14.76 - 66.77}{14.76} = -3.52$
Two-way	Consistency	$\frac{MS_R - MS_E}{MS_R}$	$\frac{14.76 - 49.93}{14.76} = -2.38$
	Absolute agreement	$\frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}}$	$\frac{14.76 - 49.93}{14.76 + \frac{184.67 - 49.93}{8}} = -1.11$

Significance testing

The calculated intraclass correlation can be tested against a value under the null hypothesis using a standard F -test. McGraw and Wong (1996) describe these tests for the various intraclass correlations we've seen and Table 3 summarizes their work. In this table ICC is the observed intraclass correlation, while ρ_0 is the value of the intraclass correlation under the null hypothesis. That is, it's the value against which you wish to compare the observed intraclass correlation. So, replace this value with 0 to test the hypothesis that the observed ICC is greater than zero, but replace it with other values such as 0.1, 0.3 or 0.5 to test that the observed ICC is greater than known values of small, medium and large effect sizes, respectively.

Table 3

ICC for Single Scores				
Model	Interpretation	F -ratio	df_1	df_2

One-way	Absolute agreement	$\frac{MS_R}{MS_W} \times \frac{1 - p_0}{1 + (k - 1)p_0}$	$n - 1$	$n(k - 1)$
Two-way	Consistency	$\frac{MS_R}{MS_E} \times \frac{1 - p_0}{1 + (k - 1)p_0}$	$n - 1$	$(n - 1)(k - 1)$
	Absolute agreement	$\frac{MS_R}{aMS_C + bMS_E}$ <p>in which</p> $a = \frac{kp_0}{n(1 - p_0)}$ $b = 1 + \frac{kp_0(n - 1)}{n(1 - p_0)}$	$n - 1$	$\frac{(aMS_C + bMS_E)^2}{\frac{(aMS_C)^2}{k - 1} + \frac{(bMS_E)^2}{(n - 1)(k - 1)}}$
ICC for Average Scores				
One-way	Absolute agreement	$\frac{1 - p_0}{1 - ICC}$	$n - 1$	$n(k - 1)$
Two-way	Consistency	$\frac{1 - p_0}{1 - ICC}$	$n - 1$	$(n - 1)(k - 1)$
	Absolute agreement	$\frac{MS_R}{cMS_C + dMS_E}$ <p>in which</p> $c = \frac{p_0}{n(1 - p_0)}$ $d = 1 + \frac{p_0(n - 1)}{n(1 - p_0)}$	$n - 1$	$\frac{(cMS_C + dMS_E)^2}{\frac{(cMS_C)^2}{k - 1} + \frac{(dMS_E)^2}{(n - 1)(k - 1)}}$

Fixed versus random effects

I mentioned earlier on that the effect of the measure/judges can be conceptualized as a fixed or random effect. Although it makes no difference to the calculation, it does affect the interpretation. Essentially, this variable should be regarded as random when the judges or measures represent a sample of a larger population of measures or judges that could have been used. Put another way, the particular judges or measures chosen are not important and do not change the research question you're addressing. However, the effect of measures should be treated as fixed when changing one of the judges or measures would significantly affect the research question. For example, in the gambling study mentioned earlier it would make a difference if the ratings of the gambler were replaced: the fact the gamblers gave ratings was intrinsic to the research question being addressed (do gamblers give accurate information about their gambling?). However, in our example of lecturers' marks, it shouldn't make any difference if we substitute one lecturer with a different one: we can still answer the same research question (do lecturers, in general, give inconsistent marks?). In terms of interpretation, when the effect of the measures is a random factor then the results can be

generalized beyond the sample; however, when they are a fixed effect, any conclusions apply only to the sample on which the ICC is based (McGraw & Wong, 1996).

References

- Eley, T. C., & Stevenson, J. (1999). Using genetic analyses to clarify the distinction between depressive and anxious symptoms in children. *Journal of Abnormal Child Psychology*, 27(2), 105–114.
- Hodgins, D. C., & Makarchuk, K. (2003). Trusting problem gamblers: reliability and validity of self-reported gambling behavior. *Psychology of Addictive Behaviors*, 17(3), 244–248.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30–46.
- Shrout, P. E. F., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing reliability. *Psychological Bulletin*, 86, 420–428.

Please, Sir, can I have some more ... centring?

We'll use the **Cosmetic Surgery.sav** data to illustrate the two types of centring discussed in the book chapter. Load this file into SPSS. Let's assume that we want to centre the variable **BDI**.

Grand mean centring



Grand mean centring is really easy since we can simply use the *compute* command that we encountered in the book. First, we need to find out the mean score for BDI. We can do this using some simple descriptive statistics. Chose **Analyze** **Descriptive Statistics** **►** **Descriptives...** to access the dialog box in Figure 1.

Select BDI and drag it to the box labelled **Variable(s)**, then click on **Options...** and select only the mean (we don't need any other information).

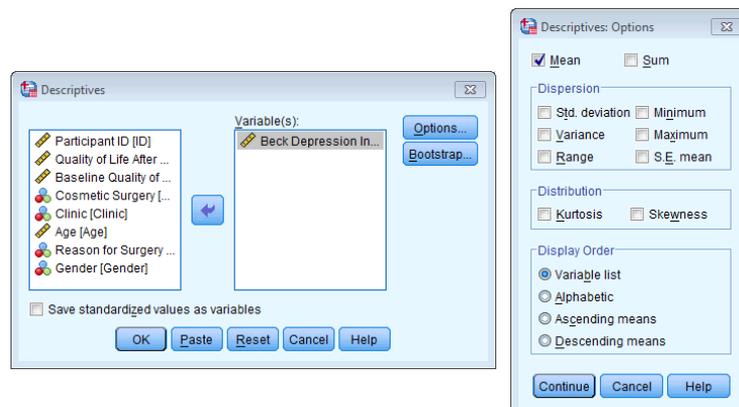


Figure 1

The resulting output tells us that the mean is 23.05:

	N	Mean
Beck Depression Inventory	276	23.05
Valid N (listwise)	276	

We use this value to centre the variable. Access the *compute* command by selecting **Transform** > **Compute Variable...**. In the resulting dialog box enter the name **BDI_Centred** into the box labelled *Target Variable* and then click on **Type & Label...** and give the variable a more descriptive name if you want to. Select the variable **BDI** and drag it across to the area labelled *Numeric Expression*, then click on **-** and then type the value of the mean (23.05). The completed dialog box is shown in Figure 2.

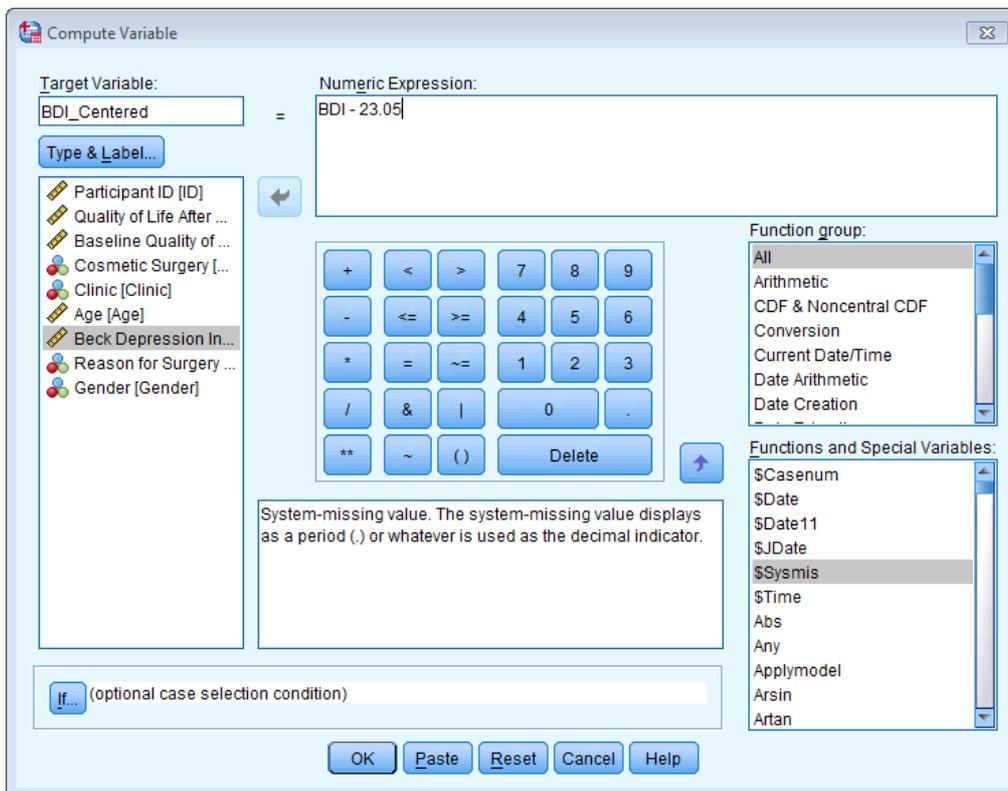


Figure 2

Click on **OK** and a new variable will be created called **BDI_Centred**, which is centred around the mean of BDI. The mean of this new variable should be approximately 0: run some descriptive statistics to see that this is true.

You can do the same thing in a syntax window by typing:

COMPUTE BDI_Centred = BDI-23.05.

EXECUTE.

Group mean centring

Group mean centring is considerably more complicated. The first step is to create a file containing the means of the groups. Let's try this again for the BDI scores. We want to centre this variable across the level 2 variable of **Clinic**. We first need to know the mean BDI in each group and to save that information in a form that SPSS can use later on. To do this we need to use the *aggregate* command, which is not discussed in the book. To access the main dialog box select **Data > Aggregate...**. In this dialog box (Figure 3) we want to select **Clinic** and drag it to the area labelled *Break Variable(s)*. This will mean that the variable **Clinic** is used to split up the data file (in other words, when the mean is computed it will do it for each clinic separately). We then need to select **BDI** and drag it to the area labelled *Summaries of Variable(s)*. You'll notice that once this variable is selected the default is that SPSS will create a new variable called **BDI_mean**, which is the mean of **BDI** (split by **Clinic**, obviously). We need to save this information in a file that we can access later on, so select

Write a new data file containing only the aggregated variables. By default, SPSS will save the file with the name **aggr.sav** in your default directory. If you would like to save it elsewhere or under a different name then click on **File...** to open a normal file system dialog box where you can name the file and navigate to a directory that you'd like to save it in. Click on **OK** to create this new file.

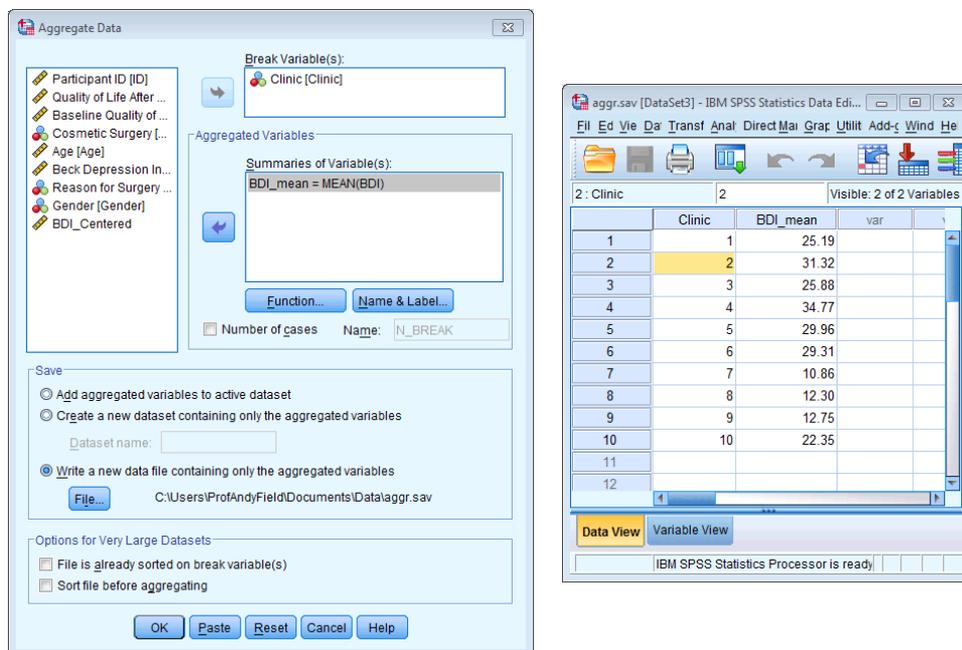


Figure 3

If you open the resulting data file (you don't need to, but it will give you an idea of what it contains) you will see that it simply contains two columns, one with a number specifying the clinic from which the data came (there were 10 clinics) and the second containing the mean BDI score within each clinic (Figure 3, right).

When SPSS creates the aggregated data file it orders the clinics from lowest to highest (regardless of what order they are in the data set). Therefore, to make our working data file match this aggregated file, we need to make sure that all of the data from the various clinics are ordered too from clinic 1 up to clinic 10. This is easily done by using the *sort cases* command. (Actually our data are already ordered in this way, but because your data might not always be, we'll go through the motions anyway.) To access the *sort cases* command select **Data** **Sort Cases...**. The resulting dialog box is shown in Figure 4. Select the variable that you want to sort the file by (in this case **Clinic**) and drag it to the area labelled **Sort by** (or click on ). You can choose to order the file in ascending order (clinic 1 to clinic 10), which is what we need to do here, or descending order (clinic 10 to clinic 1). Click on **OK** to sort the file.

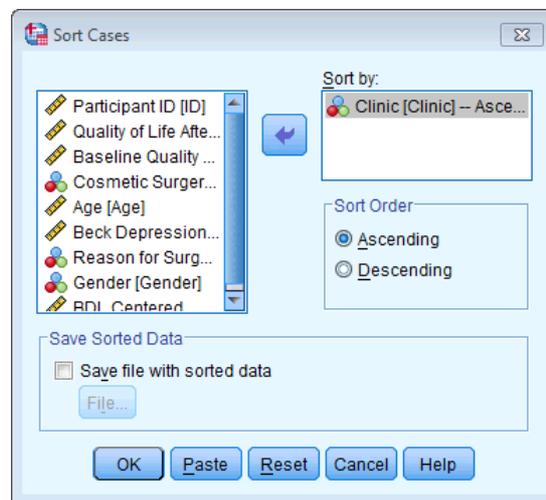


Figure 4

The next step is to use these clinic means in the aggregated file to centre the BDI variable in our main file. To do this we need to use the *match files* command, which can be accessed by selecting **Data Merge Files** **Add Variables...**. This will open a dialog box (Figure 5) that lists all of open data files (in my case I had none open apart from the one that I was working from, so this space is blank) or asks you to select an SPSS data file. Click on **Browse...** and navigate to wherever you decided to store the file of aggregated values (in my case **aggr.sav**). Select this file, then click **Open** to return to the dialog box. Then click on **Continue** to move on to the next dialog box.

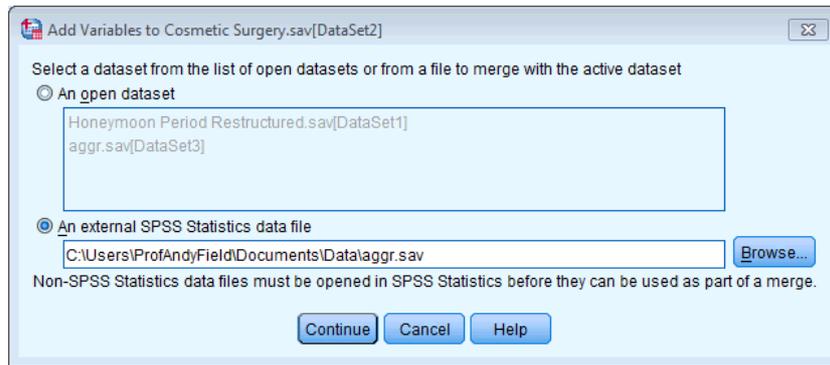


Figure 5

In the next dialog box (Figure 6) we need to match the two files, which just tells SPSS that the two files are connected. To do this click on Match cases on key variables in sorted files. Then we also need to specifically connect the files on the **Clinic** variable. To do this select Non-active dataset is keyed table, which tells SPSS that the data set that isn't active (i.e., the file of aggregated scores) should be treated as a table of values that are matched to the working data file on a key variable. We need to select what this key variable is. We want to match the files on the **Clinic** variable, so select this variable in the *Excluded Variables* list and drag it to the space labelled *Key Variables* (or click on ). Click on .

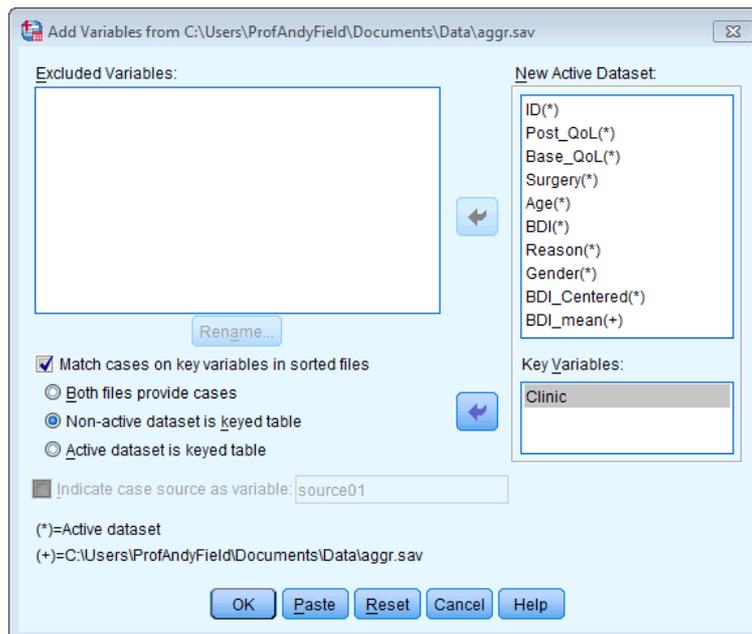


Figure 6

The data editor should now include a new variable, **BDI_mean**, which contains the values from our file **aggr.sav**. Basically, SPSS has matched the files for the clinic variable, so that the values in **BDI_mean** correspond to the mean value for the various clinics. So, when the clinic variable is 1, **BDI_mean** has been set as 25.19, but when clinic is 2, **BDI_mean** is set to 31.32. We can use these values in the *compute* command again to centre BDI. Access the *compute* command by selecting **Transform**  **Compute Variable...**. In the resulting dialog box (Figure 7) enter

the name **BDI_Group_Centred** into the box labelled *Target Variable* and then click on **Type & Label...** and give the variable a more descriptive name if you want to. Select the variable **BDI** and drag it across to the area labelled *Numeric Expression*, then click on **-** and then either type 'BDI_mean' or select this variable and drag it to the box labelled *Target Variable*. Click on **OK** and a new variable will be created containing the group centred means.

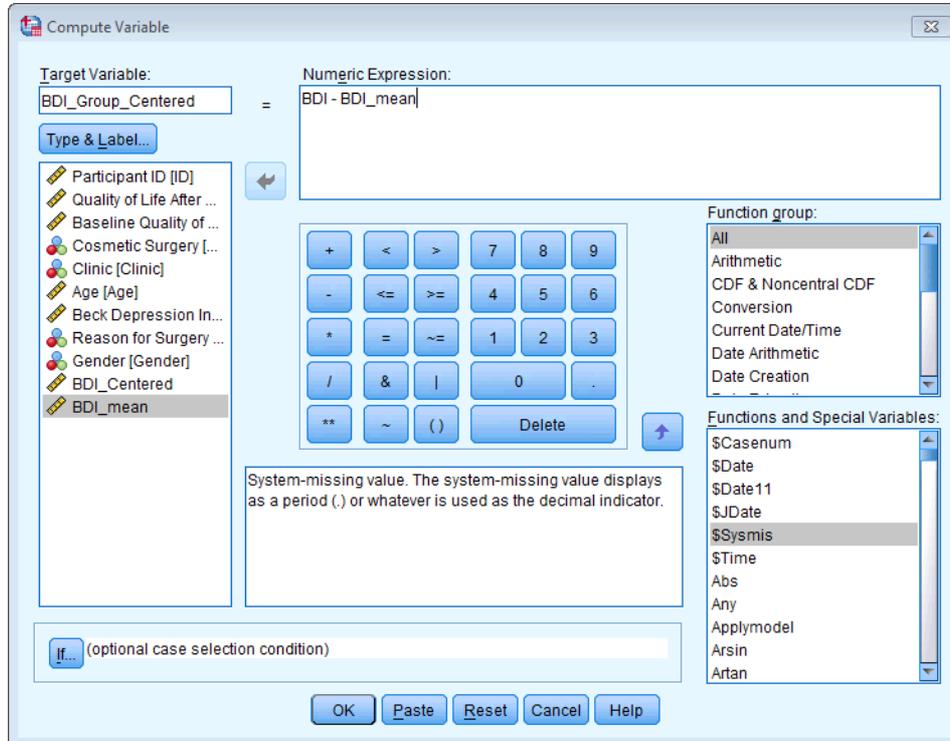


Figure 7

Alternatively, you can do this all with the following syntax:

AGGREGATE

```
/OUTFILE='C:\Users\Dr. Andy Field\Documents\Academic\Data\aggr.sav'
```

```
/BREAK=Clinic
```

```
/BDI_mean=MEAN(BDI).
```

```
SORT CASES BY Clinic(A).
```

```
MATCH FILES /FILE=*
```

```
/TABLE='C:\Users\Dr. Andy Field\Documents\Academic\Data\aggr.sav'
```

```
/BY Clinic.
```

```
EXECUTE.
```

```
COMPUTE BDI_Group_Centred=BDI - BDI_mean.
```

```
EXECUTE.
```

Please Sir, can I have some more ... restructuring?



To access the Restructure Data Wizard select **Data**  **Restructure...**. The steps in the wizard are shown in Figure 8. In the first dialog box you need to say whether you are converting variables to cases, or cases to variables. We have different levels of time in different columns (variables), and we want them to be in different rows (cases), so we need to select **Restructure selected variables into cases**. Click on **Next** to move to the next dialog box. This dialog box asks you whether you are creating just one new variable in your new data file from different columns in the old data file, or whether you want to create more than one new variable. In our case we are going to create one variable representing life satisfaction; therefore, select **One (for example, w1, w2, and w3)**. When you have done this click on **Next**. The next dialog box is crucial because it's where you set up the new data file. By default, SPSS creates a variable in your new data file called **id** which tells you from which person the data came (i.e., which row of the original data file). It does this by using the case numbers in the original data file. This default is fine, but if you want to change it (or the name **id**) then go to the section labelled *Case Group Identification* and change **Use case number** to be **Use selected variable** and then select a variable from your data file to act as a label in the new data file. I have chosen the variable **Person** from the original data set to identify participants in the new data file.

In the section labelled *Variables to be Transposed* there is a drop-down list labelled *Target Variable* which should contain an item labelled (rather unimaginatively) *trans1*. There is one item because we specified that we wanted one new variable in the previous dialog box (if we had asked for more than one new variable this drop-down list would contain as many items as variables that we requested). We can change the name of *trans1* by selecting the variables in the drop-down list and then editing their names. I suggest that you rename the variable *Life_Satisfaction*. We then need to tell SPSS which columns are associated with these two variables. Select from the list labelled *Variables in the Current File* the four variables that represent the different time points at which life satisfaction was measured (**Satisfaction_Baseline**, **Satisfaction_6_Months**, **Satisfaction_12_Months**, **Satisfaction_18_Months**). If you hold down the *Ctrl* key then you can select all four variables and either drag them across or click on . It's important that you select the variables in the correct order: SPSS assumes that the first variable that it encounters is the first level of the

repeated measure, and the second variable is the second level and so on. Once the variables are transferred, you can reorder them by using  or  to move selected variables up or down the list. Finally, there is a space to select *Fixed Variable(s)*. Drag variables here that do not vary at level 1 of your hierarchy. In this example, this means that we can select variables that are different in different people (they vary at level 2) but have the same value at the different time points (they do not vary at level 1). The only variable that we have like this is **Gender**, which did not change over the course of the data collection, but differs across people. When you have finished click on .

The remaining dialog boxes are fairly straightforward. The next two deal with the indexing variable. SPSS creates a new variable that will tell you from which column the data originate. In our case with four time points this will mean that the variable is simply a sequence of numbers from 1 to 4. So, if the data point came from the baseline phase it will be assigned a 1, whereas if it came from the 18-month follow-up phase it will be given a 4. You can select to have a single index variable (which we need here) or not to have one, or to have several. We have restructured only one repeated-measures variable (**Time**), so we need only one index variable to represent levels of this variable. Therefore, select **One** and click on . In the next dialog box (Figure 9) you can opt either to have the index variable containing numbers (such as 1, 2, 3, 4) or to use the names at the top of the columns from which the data came. The choice is up to you. You can also change the index variables name from *Index* to something useful such as **Time** (as I have done). The default options in the remaining dialog boxes are fine for most purposes, so you can just click on  to do the restructuring. However, if you want to, you can click on  to move to a dialog box that enables you to opt to keep any variables from your original data file that you haven't explicitly specified in the previous dialog boxes. The default is to drop them, and that's fine (any variables from the original data file that you want in the new data file should probably be specified earlier on in the wizard). You can also choose to keep or discard missing data. Again, the default option to keep missing data is advisable here because multilevel models can deal with missing values. Click on  to move on to the final dialog box. This box gives you the option to restructure the data (the default) or to paste the syntax into a syntax window so that you can save it as a syntax file. If you're likely to do similar data restructuring then saving the syntax might be useful, but once you have got use to the windows it doesn't take long to restructure new data anyway. Click on  to restructure the data.

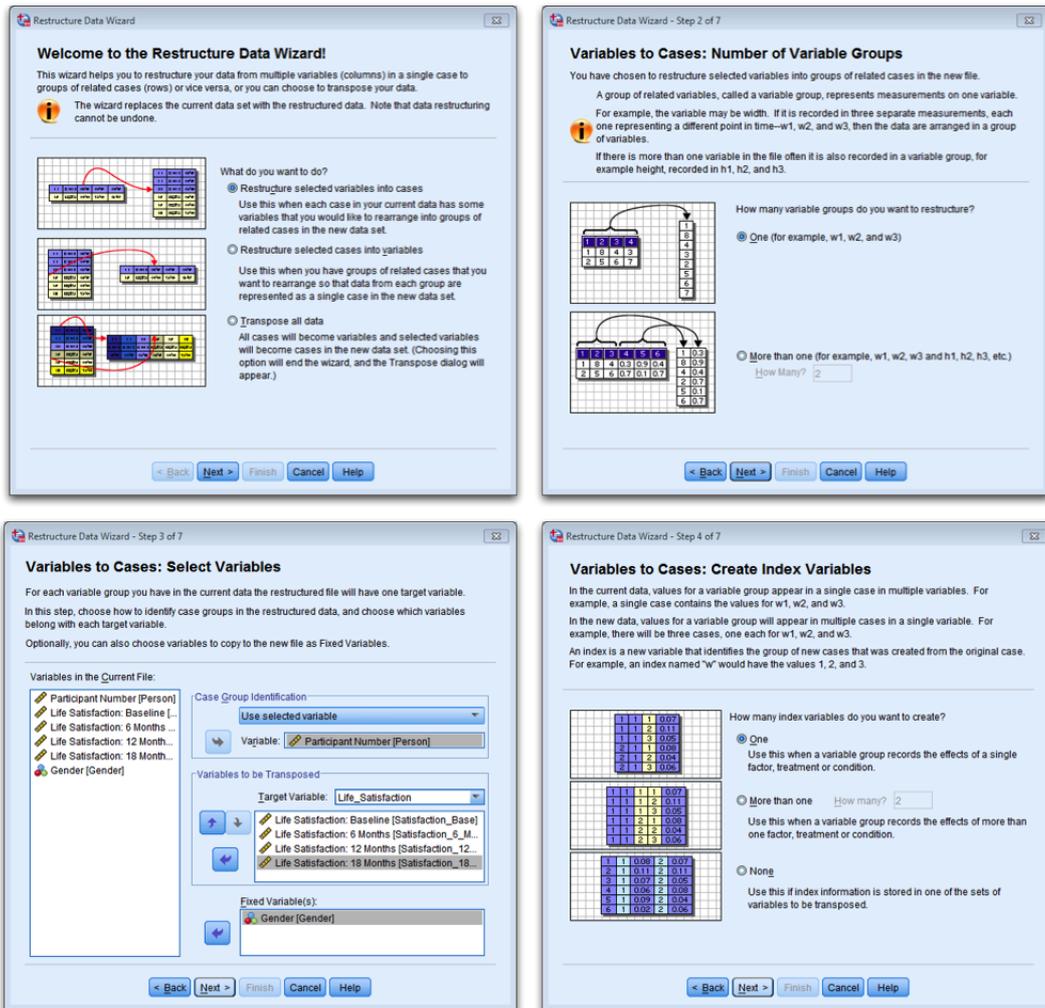


Figure 8: The Restructure Data Wizard 1

DISCOVERING STATISTICS USING SPSS

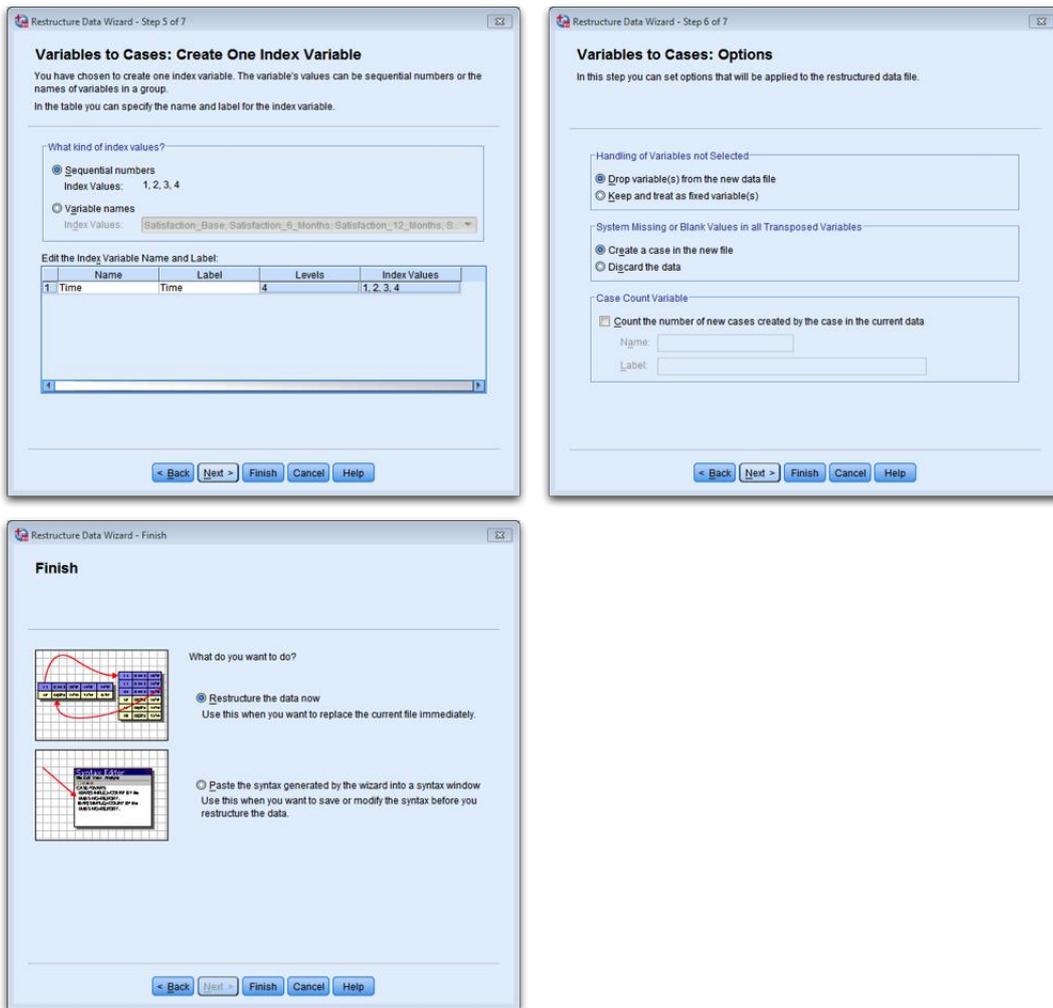


Figure 9: The Restructure Data Wizard 2