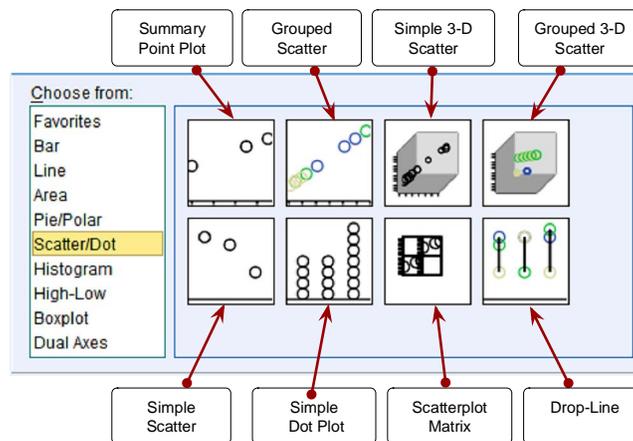


Chapter 4: Exploring data with graphs

Oliver Twisted

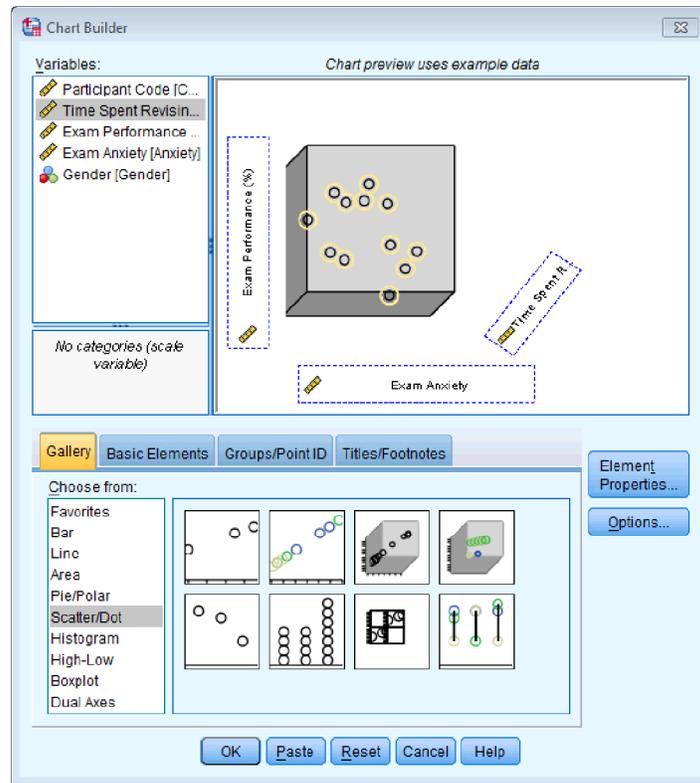
Please, Sir, can I have some more ... dimensions?

We can use the same data set as in the previous example, but this time in the Chart Builder double-click on the simple 3-D scatter icon:



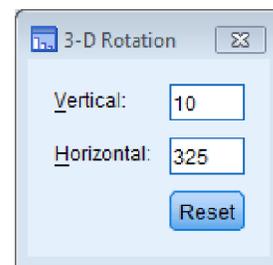
The scatter/dot gallery

The graph preview on the canvas differs from ones that we have seen before in that there is a third axis and a new drop zone (). It doesn't take a genius to work out that we simply repeat what we have done for other scatterplots by dragging another continuous variable into this new drop zone. So, select **Exam Performance (%)** from the variable list and drag it into the drop zone, and select **ExamAnxiety** and drag it into the drop zone. Now select **Time Spent Revising** in the variable list and drag it into the drop zone. Click on to produce the graph.

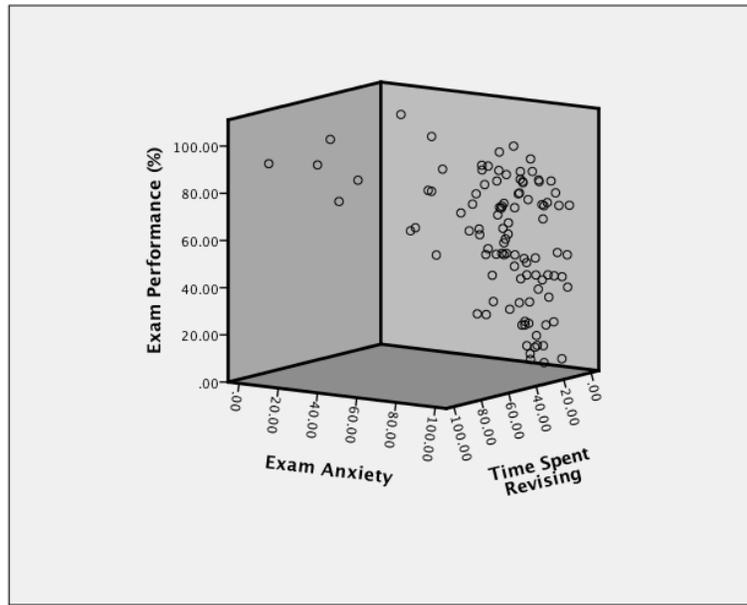


Completed Chart Builder for a 3-D scatterplot

A 3-D scatterplot is great for displaying data concisely; however, as the resulting scatterplot shows (see below), it can be quite difficult to interpret (can you really see what the relationship between exam revision and exam performance is?). As such, its usefulness in exploring data can be limited. However, you can try to improve the interpretability by changing the angle of rotation of the graph (it sometimes helps, and sometimes doesn't). To do this, double-click on the scatterplot in the SPSS Viewer to open it in the SPSS Chart Editor. Then click on  in the Chart Editor to open the *3-D Rotation* dialog box. You can change the values in this dialog box to achieve different views of data cloud. Have a play with the different values and see what happens; ultimately you'll have to decide which angle best represents the relationship for when you want to put the graph on a 2-D piece of paper! If you end up in some horrid pickle by putting random numbers in the *3-D Rotation* dialog box then simply click on  to restore the original view.



What about if we wanted to split the data cloud on our 3-D plot into different groups? This is very simple (so simple that I'm sure you can work it out for yourself); in the Chart Builder double-click on the grouped 3-D scatter icon. The graph preview on the canvas will look the same as for the simple 3-D scatterplot except that our old friend the  drop zone is back. We can simply drag our categorical variable (in this case **Gender**) into this drop zone.

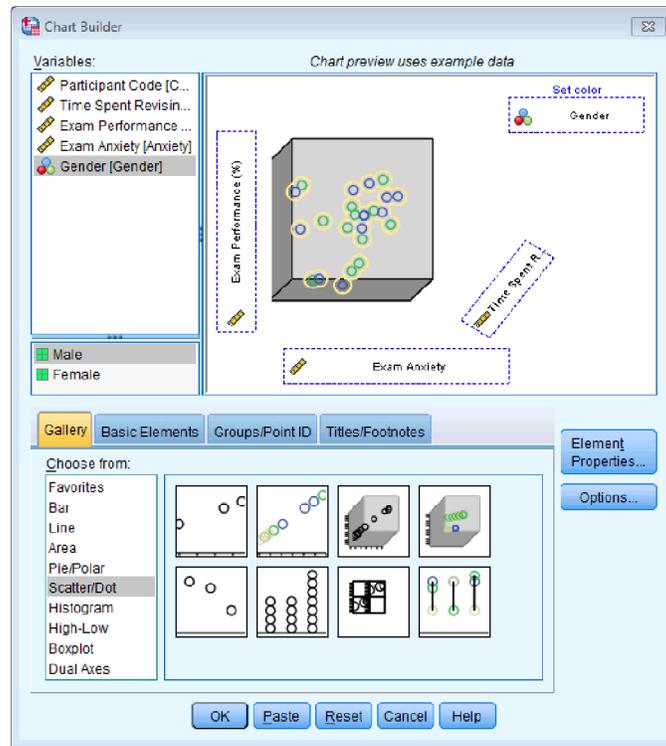


A 3-D scatterplot of exam performance plotted against exam anxiety and the amount of time spent revising for the exam



SELF TEST: Based on my minimal (and no doubt unhelpful) summary, produce a 3-D scatterplot of the same data in but with the data split by gender. To make things a bit more tricky, see if you can get SPSS to display different symbols for the two groups rather than two colours. A full guided answer is given below.

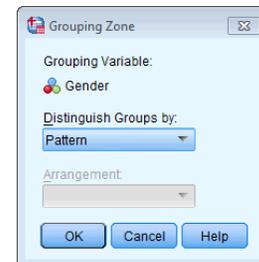
To produce this graph, first double-click on the grouped 3-D scatterplot icon in the Chart Builder (see the book for how to access the Chart Builder). The graph preview on the canvas is the same as for a simple 3-D scatterplot, except that our old friend the drop zone is back. First, we simply repeat what we have done for previous scatterplots; so, select **Exam Performance (%)** from the variable list and drag it into the drop zone, select **Exam Anxiety** and drag it into drop zone, and select **Time Spent Revising** in the variable list and drag it into drop zone. To split the data cloud by a categorical variable (in this case gender), we select this variable in the variable list and drag it into the drop zone. The completed dialog box is below.

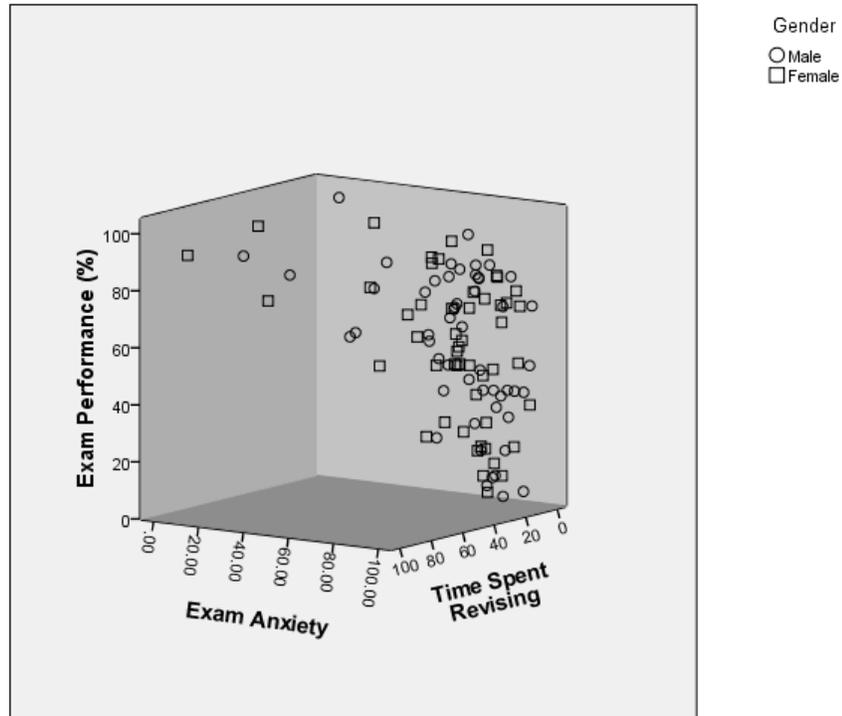


Completed dialog box for a grouped 3-D scatterplot

However, I also asked you to display the different groups as different-shaped symbols. As it stands, we have asked SPSS to produce different-coloured symbols for males and females. To change this, we need to double-click on the Set color drop zone to open a new dialog box that has a drop-down list in which *Color* will currently be selected. Click on this list to activate it and then select *Pattern*. Then click on OK to register this change. Back in

the Chart Builder the Set color drop zone will have been renamed *Set pattern*. Click on OK to plot the graph, which should look something like the one below.

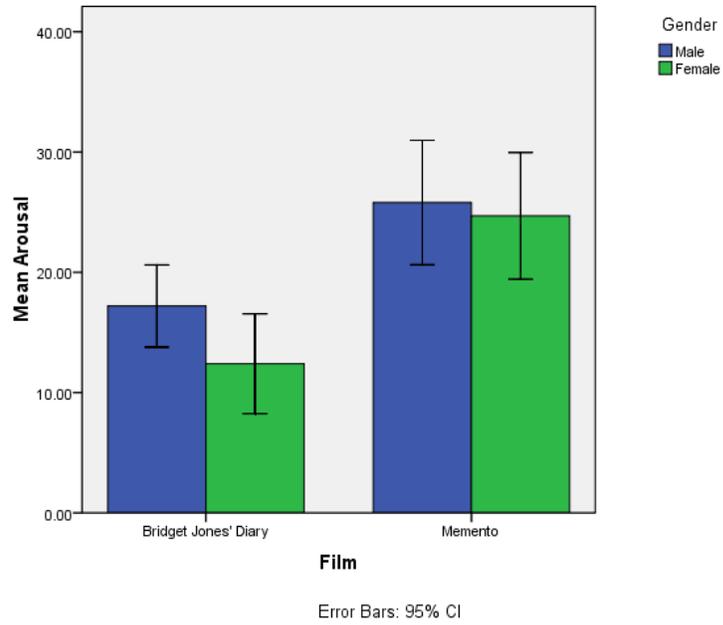




3-D grouped scatterplot

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

As an exercise to get you using some of the graph editing facilities we're going to take one of the graphs from the chapter and change some of its properties to produce a graph that follows Tufte's guidelines (minimal ink, no chartjunk, etc.). We'll use the graph for the arousal scores for the two films (*Bridget Jones's Diary* and *Memento*). The original graph looked like this:



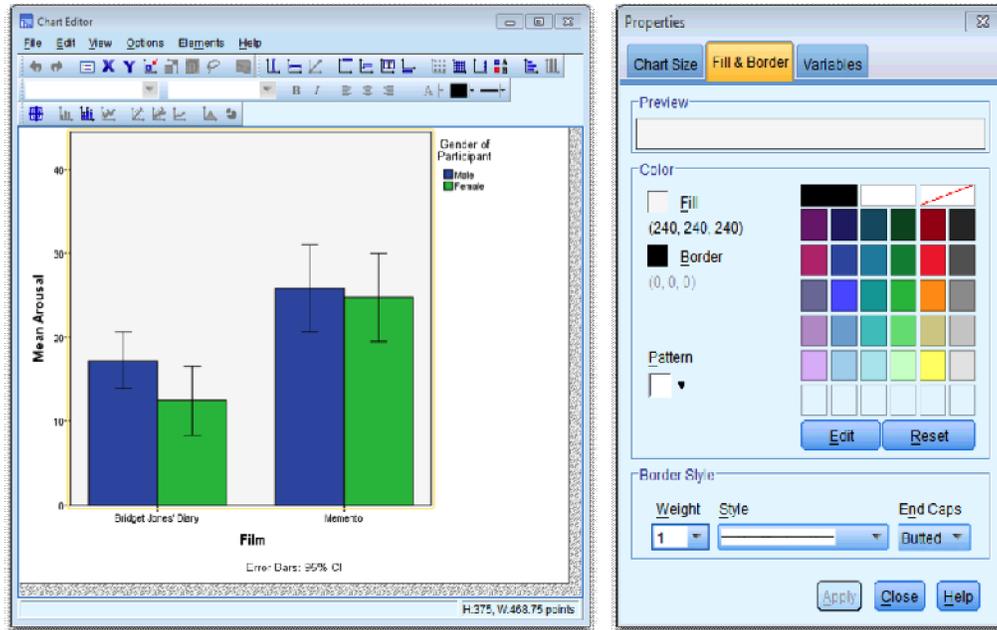
To edit this graph double-click on it in the SPSS Viewer. This will open the chart in the SPSS Chart Editor:

Double- click anywhere in the graph
to open it in the SPSS Chart Editor



Editing the chart background and border

First, let's get rid of the outside border and background colour of the graph – after all, it's just unnecessary ink! Select the border by double-clicking on it with the mouse. It will become highlighted in yellow and a *Properties* dialog box will appear:



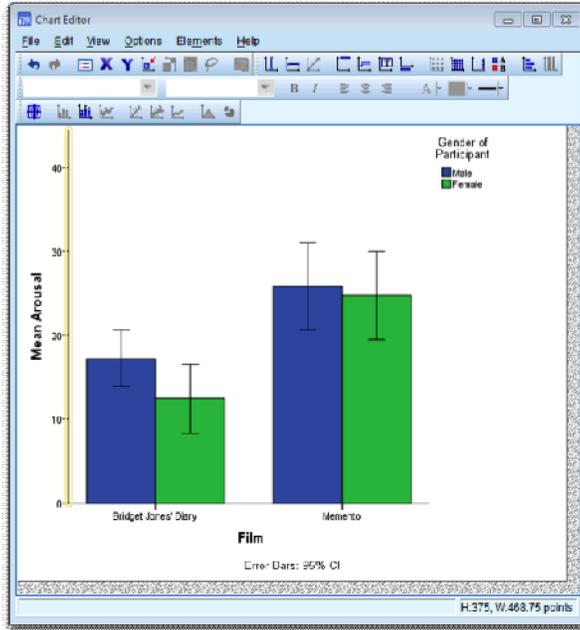
If you select the *Fill & Border* tab (as above), you can change the background and border of the chart. At the bottom there are options to change the border style, such as how thick it is (*Weight*), the style (full, dotted, etc.) and whether the lines end round, square or butted.

We can change the background and border colour by selecting a colour from the palette. To change the background colour click on *Fill* and then click on any colour from the palette. The square next to the word *Fill* will change colour. To change the border colour click on *Border* and then select a colour from the palette. Again, the square will change from black to this new colour. In this case I want us to get rid of the border and to make the background plain. Therefore, for both I want you to select 'no colour' – this is represented by . So, click on *Fill*, then click on , then click on *Border*, and then click on . To apply these changes click on ; the border and background should disappear.

Editing the axes

Now, let's get rid of the axis lines – they're just unnecessary ink too! Select the *y*-axis by double-clicking on it with the mouse. It will become highlighted in yellow and a *Properties* dialog box will appear much the same as before. This dialog box has many tabs that allow us to change aspects of the *y*-axis. We'll look at some of these in turn.

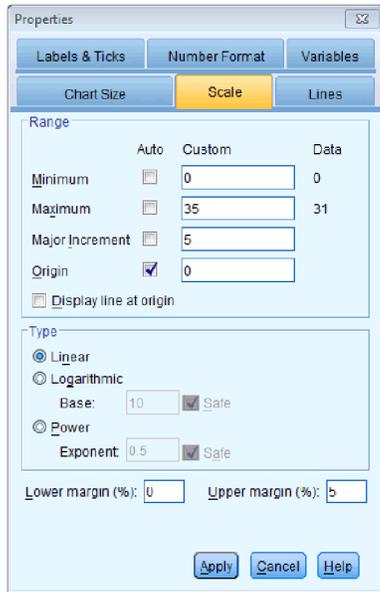
DISCOVERING STATISTICS USING SPSS



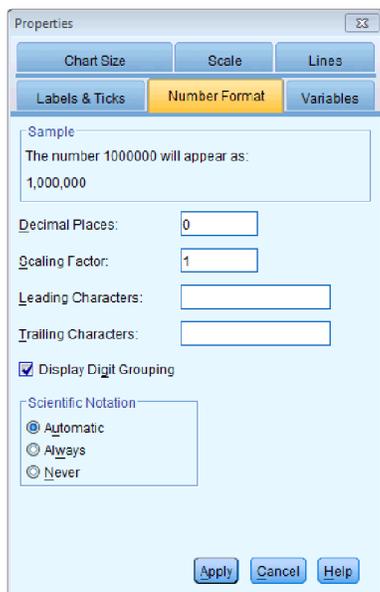
The Properties dialog box is open to the 'Scale' tab. It contains the following settings:

- Range:**
 - Minimum: Auto (0), Custom (0)
 - Maximum: Auto (40), Custom (31)
 - Major Increment: Auto (10), Custom (10)
 - Origin: Auto (0), Custom (0)
 - Display line at origin
- Type:**
 - Linear
 - Logarithmic
 - Base: 10 Safe
 - Power
 - Exponent: 0.5 Safe
- Lower margin (%): 0
- Upper margin (%): 10

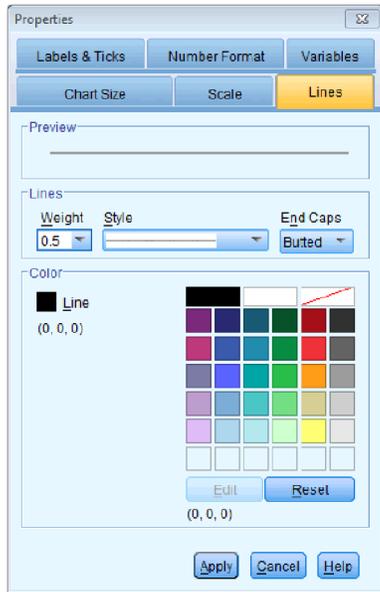
Buttons: Apply, Close, Help



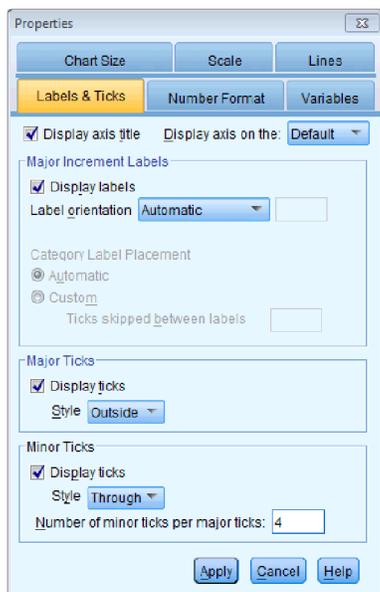
The *Scale* tab. This tab allows us to change the minimum, maximum and increments on the scale. Currently our graph is scaled from 0 to 40 and has a tick every 10 units (the major increment is, therefore, 10). However, there is a lot of space at the top of the graph. First switch off all of the autos and then change the *Maximum* from 40 to 35. In doing so we cannot have major increments of 10 (because 10 does not divide into 35 without a remainder). So, we need to change the *Major Increment* to a value that does divide into 35. Let's use 5. To make sure that SPSS doesn't rescale the minimum (it will do), also deselect auto for *Minimum* and make sure the value is set as 0. Click on **Apply** and the scale of the y-axis should change in the Chart Editor.



The *Number Format* tab allows us to change the number format used on the y-axis. The default is to have 2 decimal places, but because all of our ticks appear at values of whole numbers (0, 5, 10, 15, etc.) these decimal places are redundant. If we change the *Decimal Places* to 0 (see left) then we can get rid of these superfluous decimal places. Click on **Apply** and the decimal places on the y-axis should vanish in the Chart Editor.



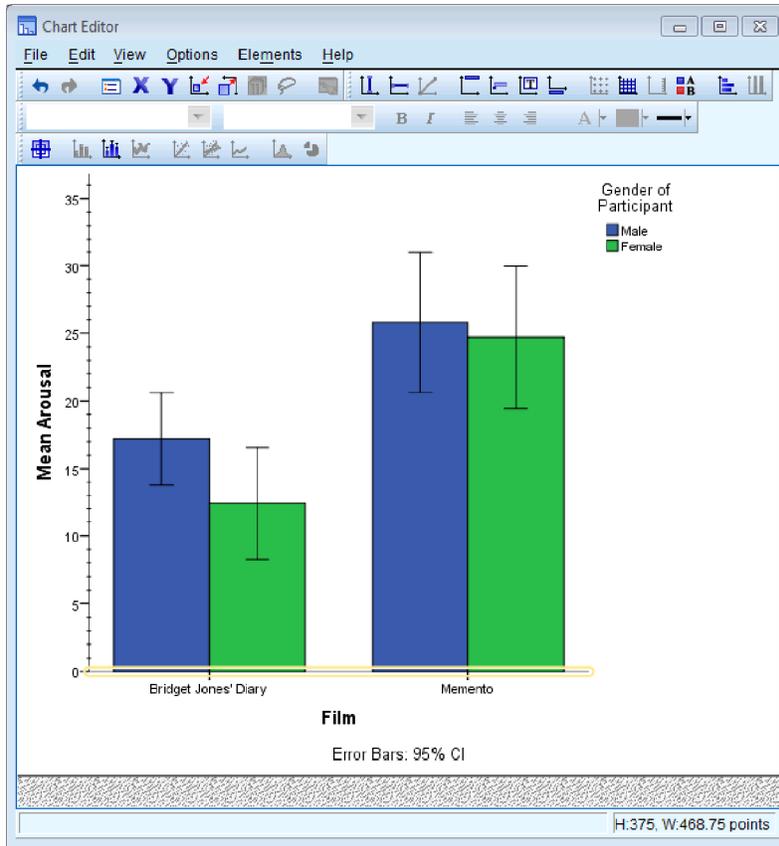
The *Lines* tab allows us to change the properties of the axis itself. We don't really need to have a line there at all, so let's get rid of it in the same way as we did for the background border. Click on  and then click on  *Line*. Click on  and the y-axis line should vanish in the Chart Editor.

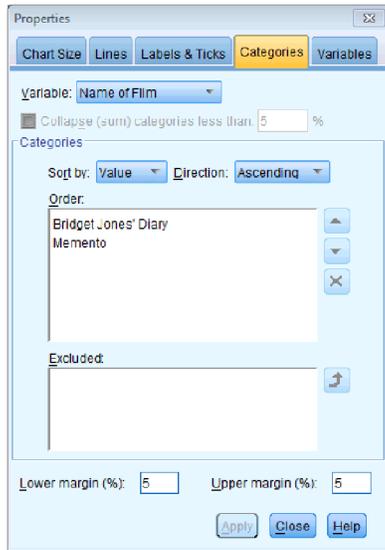


The *Labels & Ticks* tab allows us to change various aspects of the ticks on the axis. The major increment ticks are shown by default (you should leave them there), and labels for them (the numbers) are shown by default also. These numbers are important, so leave the defaults alone. You could choose to display minor ticks. Let's do this. Ask it to display the minor ticks. We have major ticks every 5, so it might be useful to have a minor tick every 1. To do this we need to set *Number of minor ticks per major tick* to be 4 (see left).

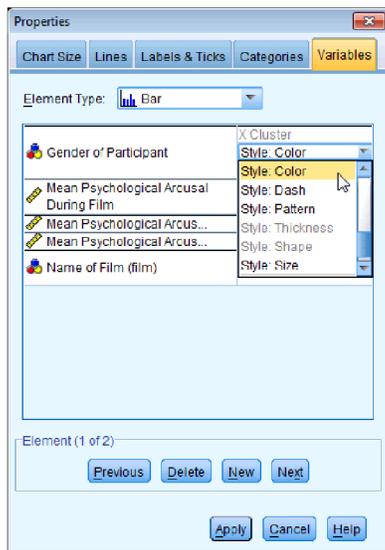
Let's now edit the x-axis. To do this double-click on it in the Chart Editor. The axis will become highlighted in yellow and the *Properties* dialog box will open. Some of the properties tabs are the same as for the y-axis so we'll just look at the ones that differ. Using what you have learnt already, set the line colour to be transparent so this it disappears (see the *Lines* tab above).

DISCOVERING STATISTICS USING SPSS





The *Categories* tab allows us to change the order of categories on this axis.

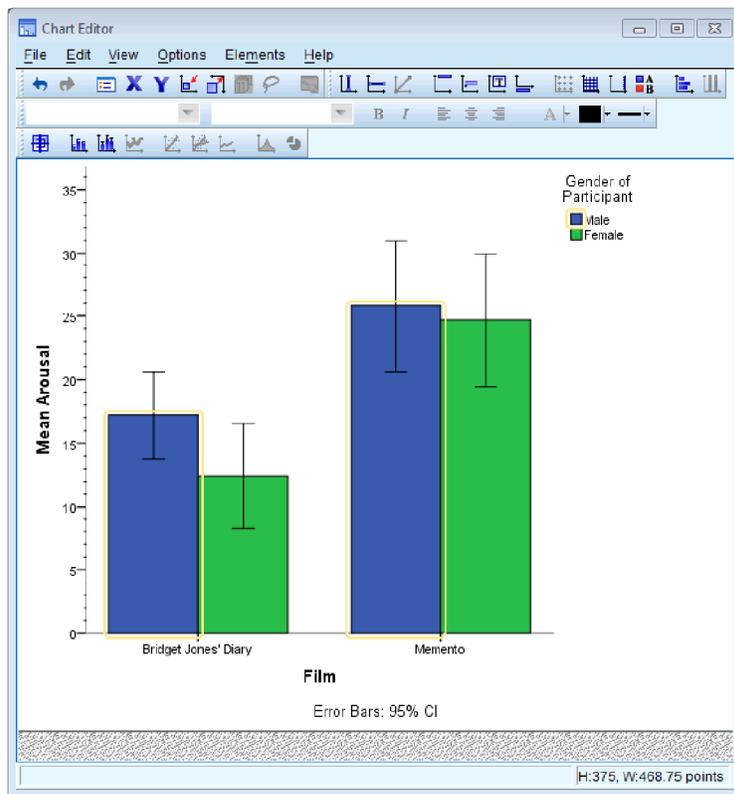
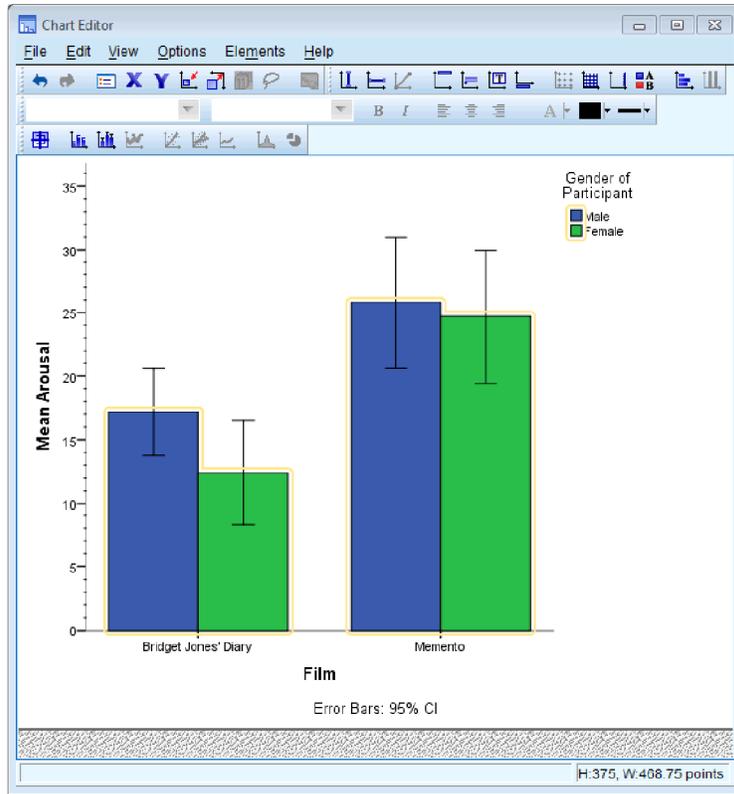


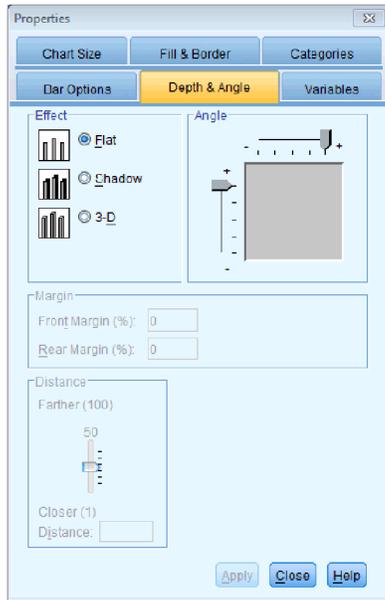
The *Variables* tab allows us to change properties of the variables. For one thing, if you don't want a bar chart then there is a drop-down list of alternatives from which you can choose. Also we have gender displayed by different colours, but we can change this so that genders are differentiated by other style differences (such as a pattern). See the drop-down list (left).

Editing the bars

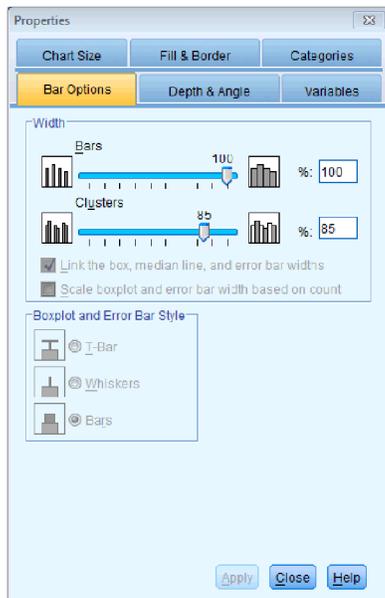
To edit the bars, double-click on any of the bars to select them. They will become highlighted with a yellow line. Let's first change the colour of the blue bars. To do this we first need to click once on the blue bars. Now instead of all of the bars being highlighted in yellow, only the blue bars will be (see below). We can then use the *Properties* dialog box to change features of these bars.

DISCOVERING STATISTICS USING SPSS



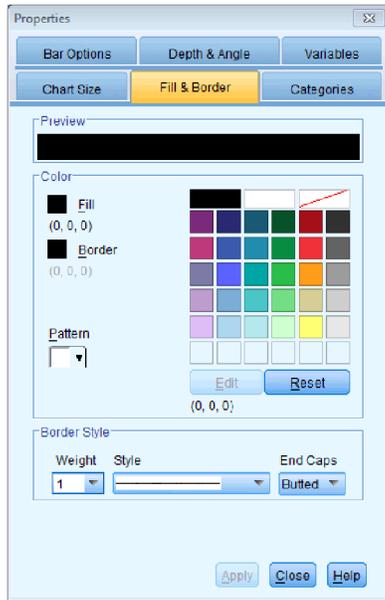


The *Depth & Angle* tab allows us to change whether the bars have a drop shadow or a 3-D effect. As I tried to stress in the book, you shouldn't add this kind of chartjunk, so you should leave your bars as flat. However, in case you want to ignore my advice, this is how you add chartjunk!



The *Bar Options* tab allows us to change the width of the bars (the default is to have bars within a cluster touching, but if you reduce the bar width below 100% then a gap will appear). You can also alter the gap between clusters. The default allows a small gap between clusters (which is sensible) but you can reduce the gap by increasing the value up to 100% (no gap between clusters) or less (a gap between clusters).

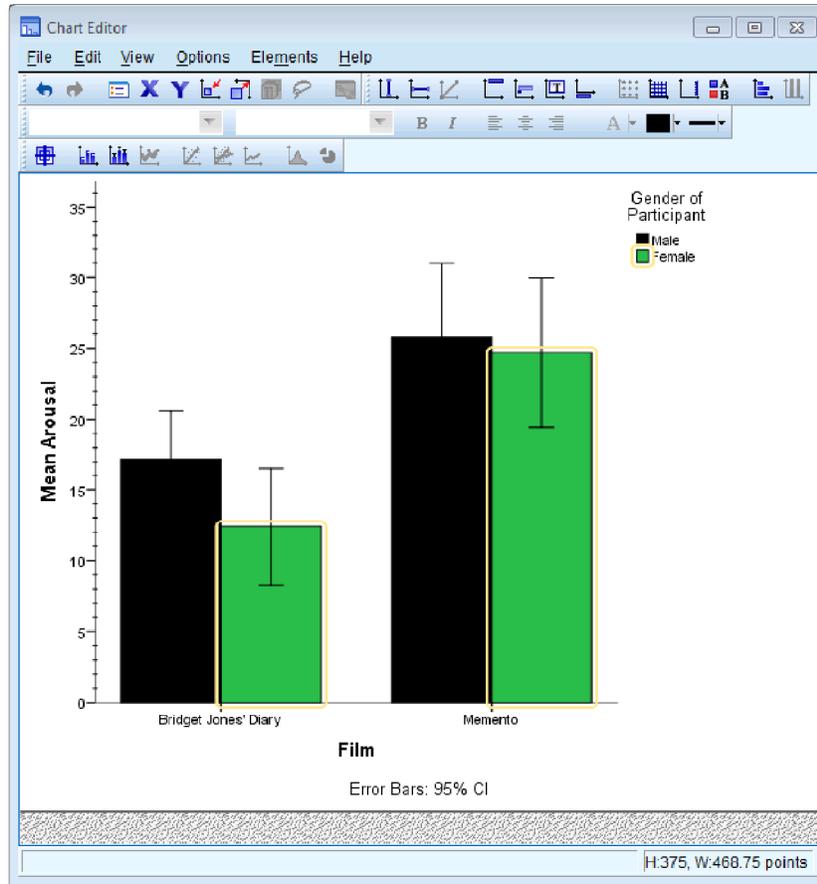
You can also select whether the bars are displayed as bars (the default) or if you want them to appear as a line (*Whiskers*) or a T-bar (*T-bar*). This kind of graph really looks best if you leave the bars as bars (otherwise the error bars look silly).



The *Fill & Border* tab allows us to change the colour of the bar and the style and colour of the bar's border. I want this bar to be black, so select the colour black from the palette and then click on (see left). Click on and the blue bars should turn black.

Now we will change the colour of the green bars. To do this we first need to click once on the green bars. Now only the blue bars will be highlighted in yellow (see below). We can then use the *Properties* dialog boxes described above to change the colour of these bars. I want you to colour these bars grey.

DISCOVERING STATISTICS USING SPSS

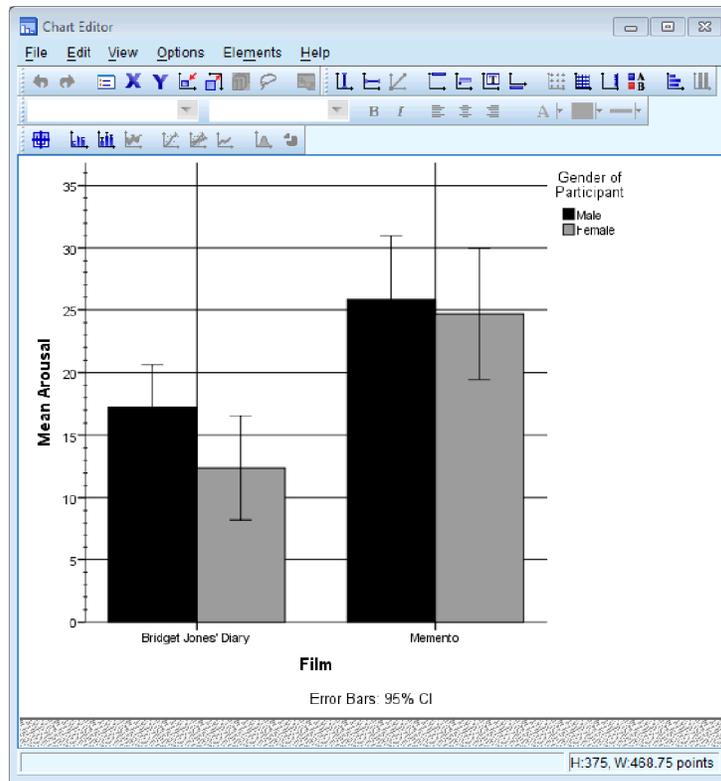




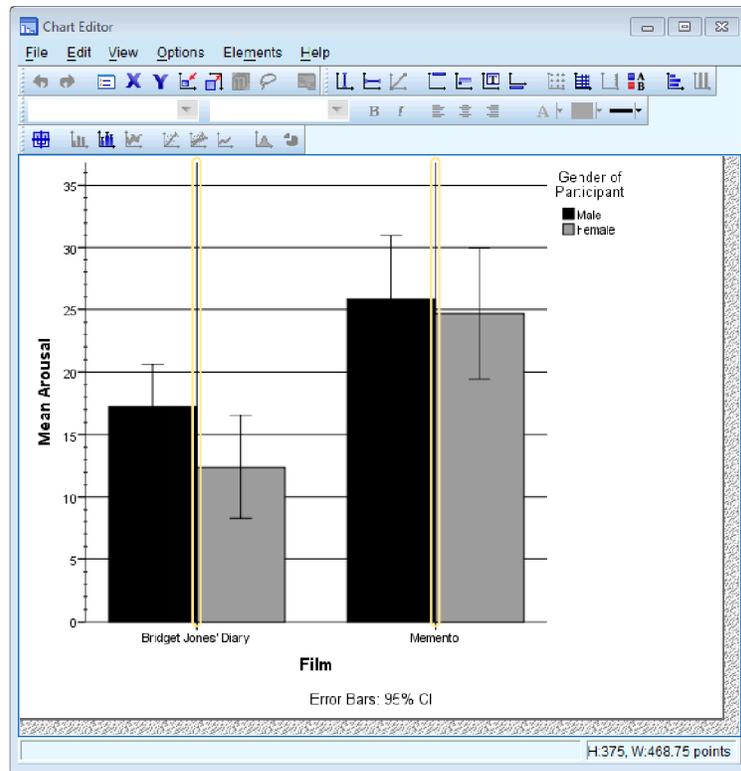
Adding grid lines

You can add grid lines to a graph simply by clicking on  in the Chart Editor. If you do this you will see that, by default, SPSS adds some rather hideous-looking lines to your graph:

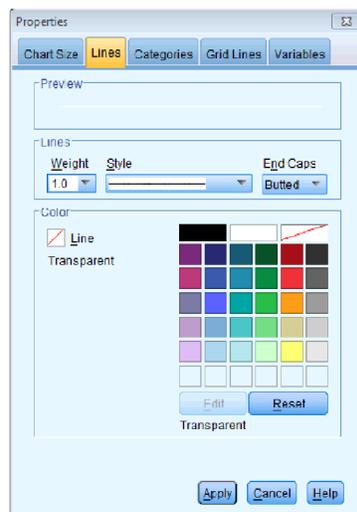
DISCOVERING STATISTICS USING SPSS



First off, we don't really want grid lines on our x-axis (the vertical grid lines), so let's get rid of them. To do this select them so that they are highlighted in yellow in the Chart Editor:



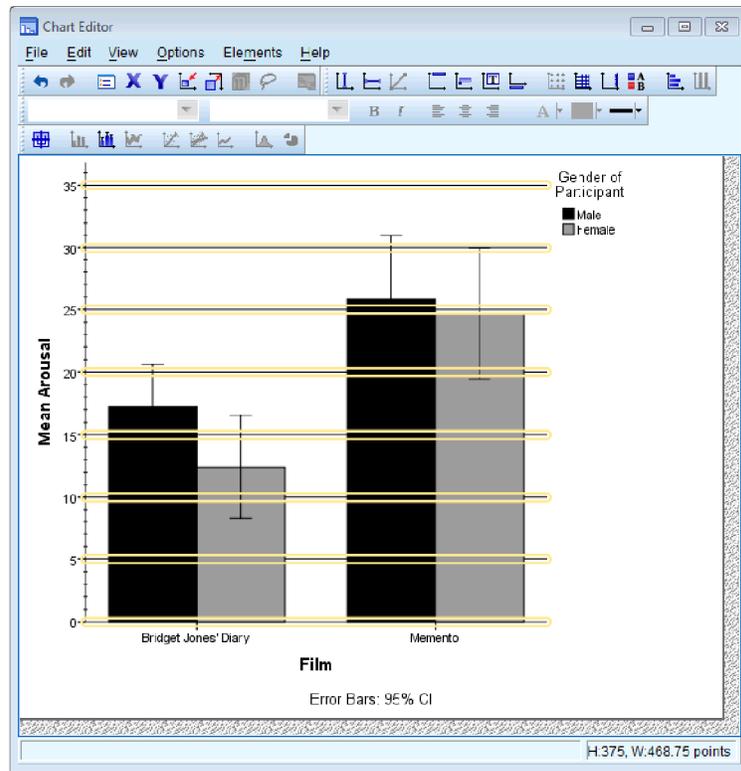
Then in the *Properties* dialog box we change the colour of these lines to be transparent (as we have done with the axis lines above):



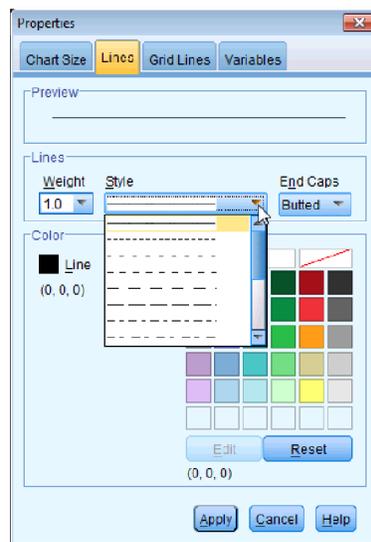
Click on **Apply** and the vertical grid lines should vanish from the Chart Editor.

Now let's edit the horizontal grid lines. To do this click on any one of the horizontal grid lines in the Chart Editor so that they become highlighted in yellow:

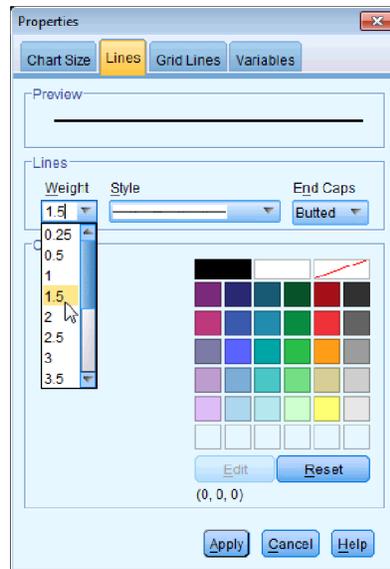
DISCOVERING STATISTICS USING SPSS



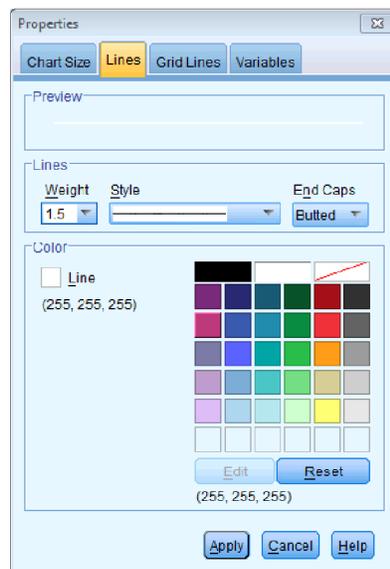
In the *Properties* dialog box select the *Lines* tab. You could change the grid lines to be dotted by selecting a dotted line from the *Style* drop-down list, but don't, leave them as solid:



Next, let's make the grid lines a bit thicker by selecting 1.5 from the Weight drop-down list:



Finally, let's change their colour from black to white. We've used the colour palette a few times now so you should be able to do this without any help (just click on the white square):

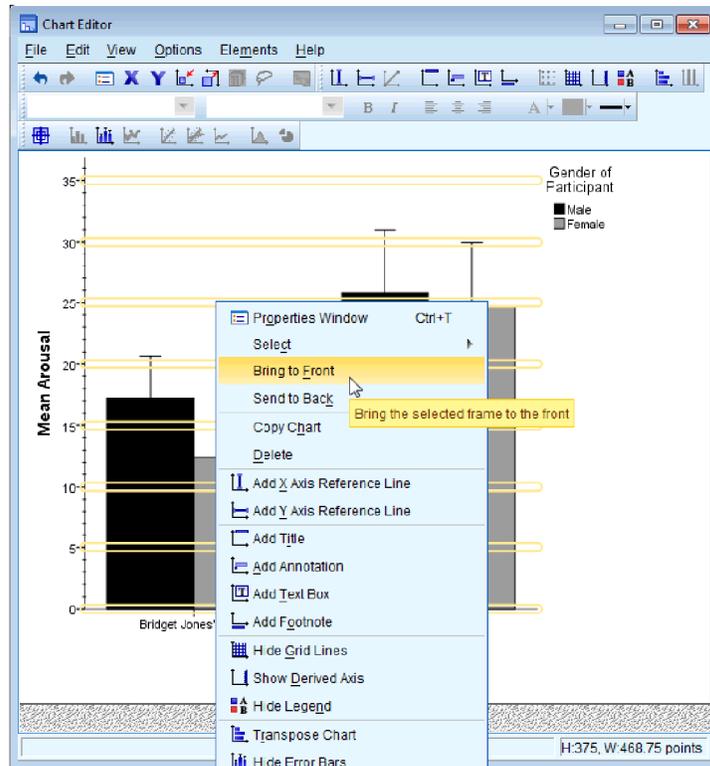


Click on [Apply](#) and the horizontal grid lines should become dotted, white and thicker. Speaking of thick, you've probably noticed that you can no longer see them because we changed the colour to white and they are displayed on a white background. You're probably also thinking that I must be some kind of idiot for telling you to do that. You're probably right, but bear with me — there is method to the madness inside my rotting breadcrumb of a brain.

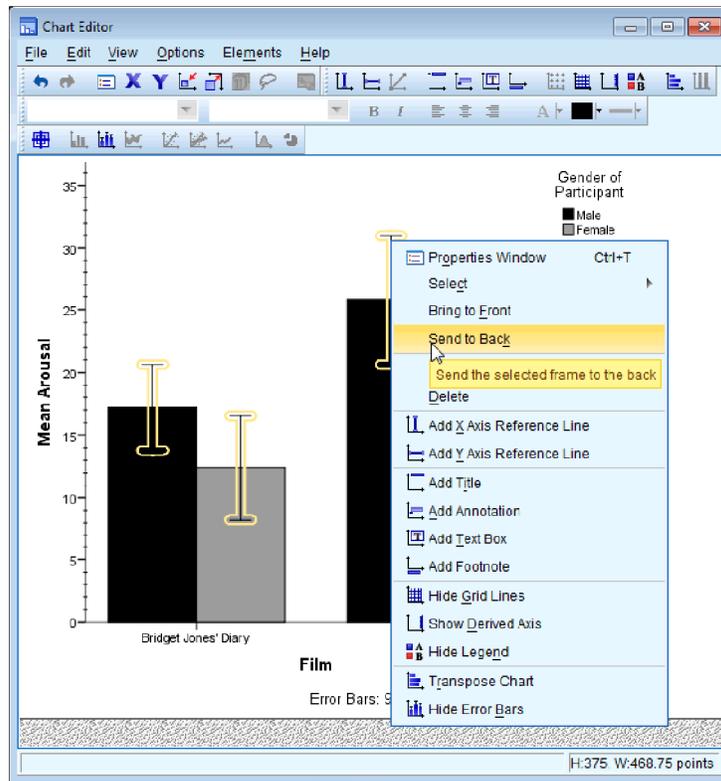
Changing the order of elements of a graph

We've got white grid lines and we can't see them. That's a bit pointless isn't it? However, we would be able to see them if they were in front of the bars. We can make this happen by

again selecting the horizontal grid lines so that they are highlighted in yellow; then if we click on one of them with the *right* mouse button a menu appears on which we can select *Bring to Front*. Select this option and, wow, the grid lines become visible on the bars themselves: pretty cool, I think you'll agree.

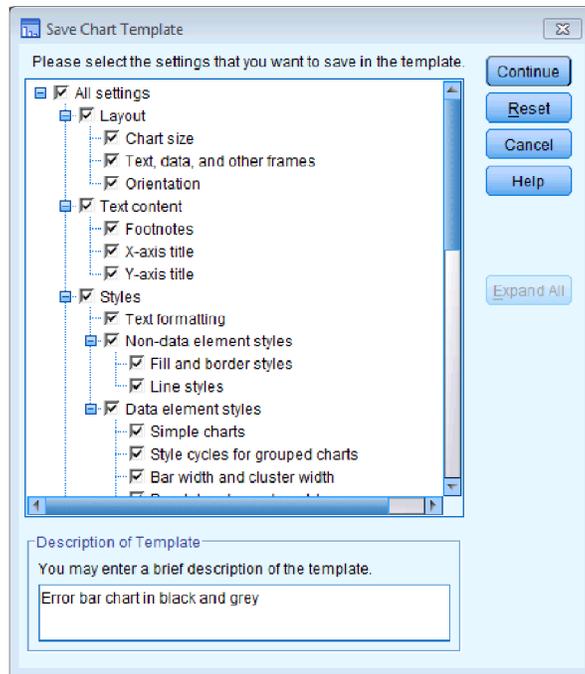


However, we still have a problem in that our error bars can be seen on top of the grey bars but not on top of the black bars. This looks a bit odd; it would be better if we could see them only poking out of the top on both bars. To do this, click on one of the error bars so that they become highlighted in yellow. Then if we click on one of them with the *right* mouse button a menu appears on which we can select *Send to Back*. Select this option and the error bars move behind the bars (therefore we can only see the top half).



Saving a chart template

You've done all of this hard work. What if you want to produce a similar-looking graph? Well, you can save these settings as a template. A template is just a file that contains a set of instructions telling SPSS how to format a graph (e.g., you want grid lines, you want the axes to be transparent, you want the bars to be coloured black and grey, and so on). To do this, in the Chart Editor go to the *File* menu and select *Save Chart Template*. You will get a dialog box, and you should select what parts of the formatting you want to save (and add a description also). Although it is tempting to just click on 'save all', this isn't wise because, for example, when we rescaled the y-axis we asked for a range of 0–35, and this is unlikely to be a sensible range for other graphs, so this is one aspect of the formatting that we would not want to save.



Click on **Continue** and then type a name for your template (I've chosen *Tufte Bar.sgt*). By default SPSS saves the templates in a folder called 'Looks', but you can save it elsewhere if you like. I have saved it in 'Documents' which is in my folder 'ProfAndyField'. Assuming you have saved a chart template, you can apply it when you run a new graph in the Chart Editor by opening the *Options* dialog box, clicking on **Add** and then browsing your computer for your template file:

