

# Bar Graphs with Two Grouping Variables.

The following techniques are demonstrated in this article:

- Graphing data organized by two grouping variables
- Reformatting those graphs to change appearance, arrangement, order, spacing, and direction of bars and to modify error bars

There are three methods for making a bar graph in Prism, each using a different graph type. Most of this article is concerned with the method that you must use to graph data organized into two grouping variables (although you may apply the method to data organized by one grouping variable). Prism uses the term *bar graph* to refer specifically to a graph produced by this method. You may be interested in the following related Step-by-Step Examples:

- When your data has only one grouping variable, the preferred method is to create a *column graph* on which data are depicted using bars (as opposed to columns of point symbols or box-and-whiskers plots). See the Step-by-Step Example "Bar Graphs with One Grouping Variable".
- You may wish to use a variation of the *XY graph* on which point symbols are replaced with "spikes" widened so as to appear as bars. This method is discussed later in this article (see page 9) and in the article "Combining a Bar Graph with a Line Graph".
- Histograms—bar graphs that illustrate frequency distributions—are covered in the article "Histograms and Frequency Distributions".

## **Creating Bar Graphs**

When you launch Prism, the Welcome dialog appears. Select **Create a new project** and indicate that you will create the initial data table and graph by choosing the **Type of graph**. Suppose we wish to create the following graph:

<sup>&</sup>lt;sup>1</sup> Adapted from: Miller, J.R., *GraphPad Prism Version 4.0 Step-by-Step Examples*, GraphPad Software Inc., San Diego CA, 2003. *Step-by-Step Examples* is one of four manuals included with Prism 4. All are available for download as PDF files at <u>www.graphpad.com</u>. While the directions and figures match the Windows version of Prism 4, all examples can be applied to Apple Macintosh systems with little adaptation. We encourage you to print this article and read it at your computer, trying each step as you go. Before you start, use Prism's **View** menu to make sure that the Navigator and all optional toolbars are displayed on your computer.

<sup>©2003</sup> GraphPad Software, Inc. All rights reserved. GraphPad Prism is a registered trademark of GraphPad Software, Inc. Use of the software is subject to the restrictions contained in the software license agreement.



As the graph shows, our data are organized by two grouping variables:

- 1. *Treatment*, with three levels—placebo, 10 mg, and 100 mg
- 2. *Time*, with two levels—pre-treatment and post-treatment

Therefore, in the Welcome dialog, select the tab for Two grouping variables.

XY One grouping variable	Two grouping variables Survival				
Row and column categ	gories for two-way ANOVA and contingency tables.				
Selected graph: Interleaved bar, vertical					
Enter error bar info into subcolumr	ns as:				
🔿 No error bars 💿 3 🛟	replicates 🔘 Mean, Standard Deviation, N 🛛 🗸 🗸				

The dialog now displays eight choices for graph type. You can switch freely between the buttons, reading the "Selected graph" descriptions if you don't find the thumbnail illustrations clear. The upper-left button—for interleaved vertical bars—matches the organization and orientation of the bars in our intended graph.

The difference between the interleaved bars and grouped bars is potentially confusing. In the thumbnail views, white bars and black bars represent different data sets. In an "interleaved" bar graph, Prism places bars for all data in a particular *row* together, mixing data from different columns in the process. In a "grouped" bar graph, Prism places bars for all data in a particular *column* together, mixing data from different rows in the process. To help you keep this straight, remember Prism's rule—it colors/patterns bars or symbols according to data set. All data in a particular data set (column) are colored/patterned in the same way.

The interleaved arrangement is the most common. Grouping is used rarely, so avoid it unless you are sure that's what you want.

With the graph type selected, indicate that you will enter data into subcolumns as **3 replicates**. That means that you will enter your replicate measurements, and Prism will compute the means and standard deviations or standard errors automatically. Click **OK** to exit the dialog. Prism creates and displays the new table.

The table contains an X column, which will hold the levels of one grouping variable. Since we indicated in the previous dialog that we will enter data as 3 replicates, Prism subdivides each data set (A, B, C, ...) to make room

for side-by-side entry of replicates (A:Y1, A:Y2, A:Y3; B:Y1, B:Y2, B:Y3; etc.). Each data set represents a level of the other grouping variable.

Click on the default table name (drop-down list on the toolbar) and type a new name for the table. That name will be used as the title for your graph and in the names of other linked sheets, although you are free to name those sheets independently.

GP-13062	¥

Enter the values show below into your table, including column headings, maintaining the appropriate organization of grouping variable.

- Since there is an X column, each row corresponds to a different level of one of the grouping variables. The X-column heading "Treatment" identifies that grouping variable. The levels of that variable are identified by the text labels "Placebo", "10 mg", and "100 mg" in the X column.
- Similarly, each Y column (i.e., each data set A, B, C, ...) corresponds to a different level of the other grouping variable. The name of that grouping variable is not included on the table, but the levels are identified in the Y-column headings "Pre-treatment" and "Post-treatment".

	X Labels A				В		
	Treatment	Pre-Treatment			F	'ost-Treatmer	nt
	Х	A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
1	Placebo	115	135	146	180	190	185
2	10 mg	113	127	135	126	150	140
3	100 mg	120	152	160	125	135	142

Which levels should you put in rows, and which in columns? Your choices will affect labeling, grouping, and appearance of the bars. Prism will use the levels in the X column as labels under the graph baseline (identifying the groups of bars) and the levels in the Y-column headings as elements in the legend. Bars from *different* columns of the data table will be graphed with *different* color/fill patterns, while bars from different rows of the *same* column are shown, repeated, with the *same* color/fill. If you don't get the arrangement you want, you can change that without redoing your data table, as we'll see in subsequent sections (see "Rearranging Bars" on page 5 and "Transposing Data" on page 7). If your data table contains three or more data sets, it will probably be easier to choose **Change... Graph Type** and select a different graph type.

You can change the number format if you wish. Select all of the columns you want to change, then choose **Number Format...** from the drop-down list under the **Change** button.

As soon as you have entered your data, Prism creates a graph automatically. In the Navigator, choose the sheet **GP-13062 graph**.



Prism automatically graphs all the data sets (columns A and B) on the data table. You may easily add or remove data sets from the graph (click the **Change** button, then choose the **Data on Graph** tab in the **Format Graph** dialog.

Note carefully how the information on your data table is used to create the bar graph:

- Text entries in the X column are placed beneath the appropriate group of bars.
- Labels for the data sets are placed in the legend, where they identify the data sets by color/pattern. As we noted earlier, Prism plots data from the same data set using the same color and pattern. To change colors, patterns, and borders of the bars, double-click on one of the bars to open the **Format Bars** dialog.
- The X column heading becomes the initial horizontal "axis" title, although you can edit the title later.
- The name of the data table becomes the title of the graph.
- Error bars are produced automatically. Prism averages replicate values and plots the mean and error bars.

### **Formatting Bar Graphs**

The following table shows tips for common format changes:

То:	Do this:
Change bar appearance	Double-click on a bar. In the <b>Format Bars</b> dialog, choose the <b>Appearance</b> tab. Verify the data set (to change all data sets together, use the <b>All</b> button):
	Change Fill, Pattern, and Border.
Move the legend	Click on one legend element, hold the <b>Shift</b> key, and click the other element. Drag both legend elements on the graph to move them. Do fine positioning using the arrow keys.
Adjust axis title position	Click on axis title (avoiding tick labels) to display two-headed-arrow cursor. Drag title away from, or closer to, the axis.
Adjust tick label positions	Double-click on an axis. Choose the appropriate <b>axis</b> tab in the <b>Format Axes</b> dialog, then adjust the <b>Numbering/labelingDistance from</b> <b>axis</b> setting.
Edit the Y axis title	Click on the default "YTitle". The text will flip to horizontal. Type the replacement title. When done, click elsewhere to flip the title back. Use the $\alpha$ button on the tool bar to enter Greek letters.
Convert the error bars from <i>SEM</i> to <i>SD</i>	Double-click on a bar, then choose the <b>Appearance</b> tab. Press the <b>All</b> button to change all data sets at once. Set <b>Error values</b> to <b>Mean &amp; SD</b> .

# **Reorganizing Bar Graphs**

#### **Rearranging Bars**

To change how the bars are arranged, double-click on one of the bars to open the **Format Bars** dialog, then choose the **Order and Direction** tab.

pearance order and Direction	
Overtical Overtical	
Baseline	
<ul> <li>Automatic</li> <li>Bars begin a</li> </ul>	at Y=
Space between bars	
Space between adjacent bars:	50 % of column width
Additional space between groups	s of bars: 100 % of bar width
Order of data sets (left to right)	
GP-13062:Post-Treatment	Тор
	Reverse
	Down
	Bottom
Polationship with proceeding data	ant .
Interference d     Charling used	
Interleaved     Stacked	U aroupea

A dialog is displayed that allows you to change bar direction, spacing, and order. You can change the graphical relationship between data sets, choosing among **Interleaved**, **Stacked**, or **Separate** bar arrangements, to produce the following arrangements.



If your data table contains three of more data sets, it will probably be easier to choose **Change... Graph Type** and select a different arrangement from the graph thumbnails.

#### **Transposing Data**

Simply changing the relative arrangement of bars (interleaved, stacked, separate) may still not yield the graph you want. In that case, try transposing the data table. With the data table displayed, click **Analyze**, choose **Data manipulations... Transpose X and Y**, and check the option to **Create a new graph of the results**. Your original data table and graph are not modified, but the transposed table is produced on a Results sheet and a new graph of the transposed data is created. If, for example, you transform this data table...

	X Labels	Α	В	С
	X Labels	Control	1 Week	1 Month
	Х	Y	Y	Y
1	No treatment	144	138	143
2	Hypophysectomy	150	99	51

to produce this Results sheet ...

X Labels		Α	В
	X Labels	No treatment	Hypophysectomy
	Х	Y	Y
1	Control	144	150
2	1 Week	138	99
3	1 Month	143	51

...Prism assigns colors and patterns to the bars in the same way before and after transposing, but since columns and rows are interchanged, so is the color/pattern scheme.



This may be useful even when your data has only one grouping variable, as in the following example. In this instance, you can use transposition to switch between having all bars shown in the same way and having bars shown differently.



#### Adjusting Spacing between Bars: Uniform Changes

You can change the horizontal spacing of bars from the **Format Bars** dialog (under the **Order and Direction** tab).

Space between bars				
Space between adjacent bars:	15	% of c	olu	mn width
Additional space between group	os of ba	ars: 50	)	% of bar width

Here is the effect of a decrease in the gap between bars (note that the bars widen in the process):



#### Adjusting Spacing between Bars: Custom Changes

Many users wish to introduce additional spacing into a bar graph to isolate one group of bars in particular—not to be confused with the "additional space" offered in the dialog shown above, which is placed between all groups. You can create this extra space by inserting an empty row in the data table in the appropriate place. For example, inserting the empty row (2) in the following data table...

	X Labels	А			
	X Labels	Low Dose			
	Х	A:Y1	A:Y2	A:Y3	
1	Control	115	135	146	
2					
3	5 mg	113	127	135	
4	25 mg	120	152	160	

...produces this change in the graph:



Note that this works only for a regular bar graph (text in the X column), not a column bar graph (no X column).

Remembering that *Prism shows values in any given data set in the same way* suggests another way to single out one, or a few, bars to be shown differently. You can always change a bar or symbol appearance by moving data to another column (data set). Thus this table...

	X Labels	А		E	3		
	X Labels	Data Set-A		Data Set-A		Data	Set-B
	Х	Mean	SEM	Mean	SEM		
1	Control	119	8				
2	2 hrs			188	11		
3	4 hrs			141	12		
4	8 hrs			113	20		
5	Follow-up	121	5				

..leads to this graph (when the *interleaved* bar arrangement is selected):



The problem with this approach is evident if you look closely at the positioning of the bars and the alignment of bars with text labels. Prism leaves space for the "missing" bars, that is, it interprets the empty cells as if they contained zeros. A better solution may be to set up your table to make an XY graph (format X column for **Numbers**), as shown below.

	X Values	4	1	E	3
	X Title	Data Set-A		Data	Set-B
	Х	Mean	SEM	Mean	SEM
1	1.0	119	8		
2	3.0			188	11
3	4.0			141	12
4	5.0			113	20
5	7.0	121	5		

Here, the X coordinate determines the horizontal position of the symbols. The resulting graph will show your values as point symbols, but you can change those symbols to "spikes" (in the **Format Symbols and Lines** dialog, open the **Shape** list box and make the fourth choice from the bottom).

Format Symbols and Lines	
Appearance Order	
Data set: - **Data Table-2:Data Set-	3 🗸 🗸 🖌 All
Currieda	Plot: Mean & SEM 💌
Show symbols Shape:	Border color:
Color: 🗾 🗾 Size:	Border thickness: 1 pt 💌
Error bars	Ţ
Color: Dir.: Abov	<ul> <li>►</li> <li>T</li></ul>
Connecting line/curve	
Show line Style:	Start line at origin
Color: Thickness:	Pattern:
Area Fill Show fill Below	× +
Fill color: 📃 🔹 Area pa	Pattern color:
Which Y axis?	
Cert Yaxis O Right Yaxis	
	Help Cancel OK

Finally, change the numbers along the X axis to text labels using "custom ticks" (double-click on the X axis, then choose **Custom Ticks**). In the **Customize Ticks and Gridlines** dialog, choose to show **Custom ticks only**. Enter the definitions for the five custom labels—for each, fill in the **Position (X=)**, **Label**, and **Tick Style** boxes and then click **Add**. As you proceed, the definitions are listed in the box below. When you're finished, the dialog should look like this:

Customize Ticks and Gridlines			
Show © Custom ticks only O Regular ticks only O Both			
Position (X=)	Label	Tick Stule	
		None	~
1.0 3.0 4.0 5.0 7.0	Control 2 hrs 4 hrs 8 hrs Follow-up	None None None None None	
Add Change Delete			
Help	Sort	Cancel	ОК

A custom tick label can be blank space, that is, you can cover over a regular tick label with a "label" containing no text content. With the cursor in the **Label** box, press the spacebar once.

