4 Drawing charts and graphs

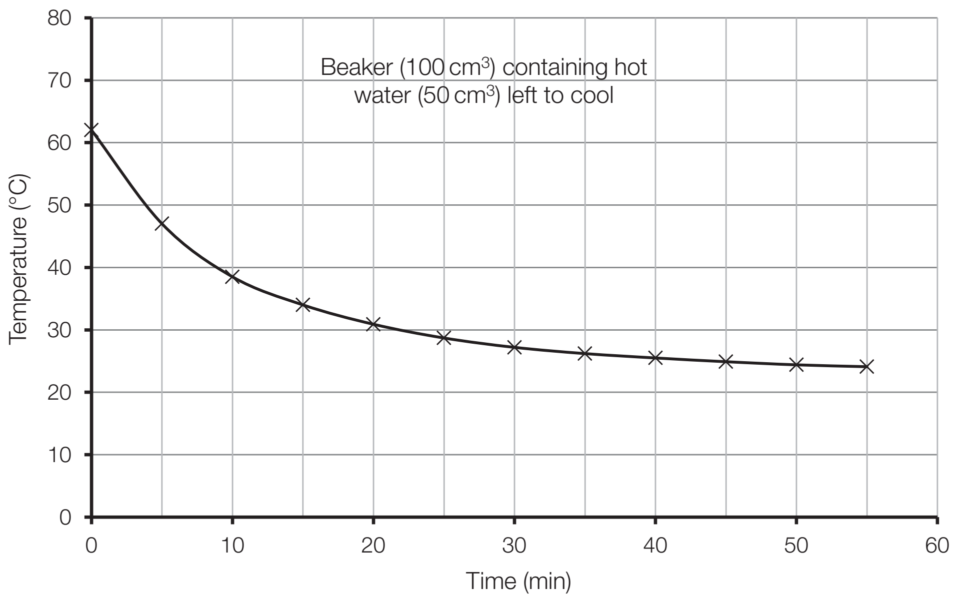
Key words: line graph, bar chart, scatter graph, independent variable, dependent variable, time series, axis, horizontal axis, vertical axis, *x*-axis, *y*-axis, origin, range, scale, tick mark, tick mark label, axis label, unit, data point, coordinate, *x*-coordinate, *y*-coordinate.

When drawing a chart or a graph, it is important to think about the purpose of doing this and what kind of display is best for representing the data. This aspect is discussed in the previous chapter (Chapter 3 *Choosing how to represent data*). This chapter focuses on the details of constructing good charts and graphs, and how to make appropriate choices when drawing them by hand on graph paper.

4.1 The important features of a chart or a graph

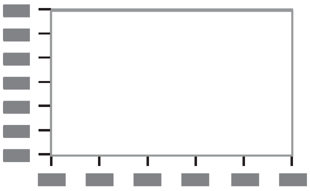
Line graphs are very common in science. Figure 4.1 shows an example of a line graph that will be used to illustrate its important features but the principles also apply to ***bar charts*** and ***scatter graphs***.

Figure 4.1 An example of a line graph

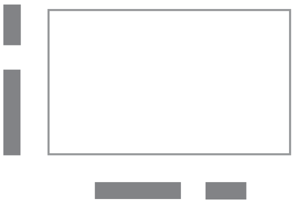


The following is a summary of the points that need to be considered during the construction of a line graph. These are discussed in more detail later in this section.

A graph has two ***axes*** drawn at right angles: the ***horizontal axis*** (or ***x-axis***) along the bottom and the ***vertical axis*** (or ***y-axis***) up the side.

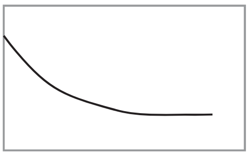


On each axis, there are ***tick marks*** (the little marks at regular intervals along each axis). There are also ***tick mark labels*** (the numbers next to the tick marks). Note: for a bar chart, one of the axes (usually the horizontal axis) would have only category labels instead.



Each axis has a label. The ***axis label*** shows the name of the variable and its unit.

The ***data points*** are the values plotted on the graph. Each point is plotted using a pair of values for the variables (the ***x-coordinate*** and the ***y-coordinate***). Note: for a bar chart, the bars would be plotted using the data values for each category.



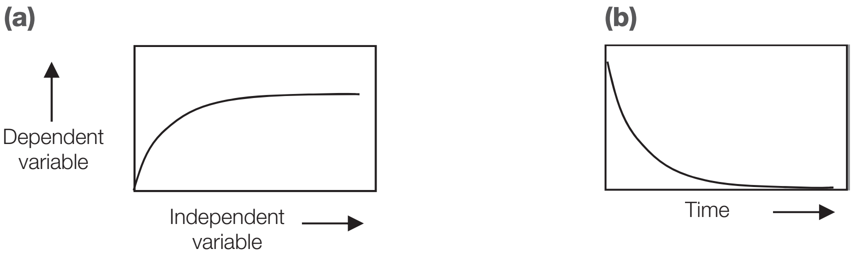
A line is drawn, which either connects all of the data points or is a line of best fit. A scatter graph may show just the data points or it may also have a line of best fit.

Finally, a graph should have a *title* that describes what the graph is showing. If there is more than one line on the graph, there will also need to be a *legend* or *key* to show what each line represents.

4.2 Choosing the axes

A graph shows the relationship between two variables. Usually, the ***independent variable*** is plotted on the ***horizontal axis*** or ***x-axis*** and the ***dependent variable*** on the ***vertical axis*** or ***y-axis*** (Figure 4.2a). Many graphs in science show how something varies over time – a ***time series*** graph. Here, time is treated as the *independent variable* (Figure 4.2b).

Figure 4.2 Choosing the axes



When plotting a graph by hand, another choice is how to orient the graph paper – landscape (wide) or portrait (tall). A landscape graph is nearly always better, especially if there is a lot of variation, as it is easier to identify the gradients. However, a portrait graph might suit rapidly increasing changes, for example exponential growth of bacteria or the rate of a chemical reaction or radioactive decay.

4.3 Choosing the range of each axis

The ***range*** of the axis refers to the lower and upper limits of the values shown on the axis. It needs to be chosen to cover the ***range*** of data. (Note that the term 'range' is here being used to refer both to the axis and to the data. Different uses of the term 'range' are explained in the glossary.) For the data shown in Figure 4.1, the values of time vary from 0 to 55 minutes and the temperature varies from about 25 °C to just over 60 °C. One question is whether to include the ***origin*** (with both axes starting at zero). Many graphs include the origin but not all do. Sometimes the data can be shown more clearly when an axis does not start at zero. For example, Figure 4.3 shows how the shape of a line can be shown more clearly by changing the range of the vertical axis.

Figure 4.3 Choosing whether to include the origin

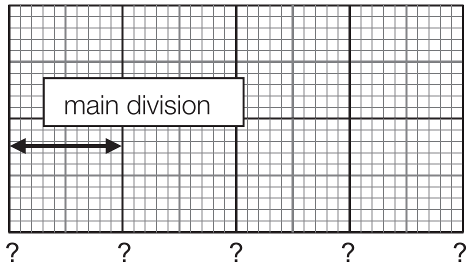
It is possible to find graphs that are similar to Figure 4.3b but which use a convention to show that the range does not start at zero. Such graphs include the zero at the start of the axis but then use a 'squiggly line' or 'zig-zag' to indicate that part of the axis has been 'cut out'. This convention is not used in scientific practice, and should be avoided in school science – the split scale can confuse pupils. When interpreting graphs, though, it is important for pupils to pay attention to the values on the scales and to know whether the graph starts at the origin or not.

Figure 4.3b shows the variation in values more clearly, but the lack of an origin can be misleading. If the *ratios of the values* are meaningful then Figure 4.3a is better for comparing sizes. The striking difference between the shapes of these lines in this example illustrates the importance of identifying the range of each axis when interpreting a line graph.

Note that in Figure 4.1 at the start of this chapter, the origin is included. Although it is not meaningful to compare the relative sizes of temperatures measured in °C, a value of 0 °C is convenient in this case for starting the vertical axis.

Sometimes, the variables plotted on the axes include negative as well as positive values. Examples of such variables include temperatures measured in °C, velocity on a velocity–time graph, or the potential difference across a component. In such cases, the origin would not be plotted at the bottom left of the graph but higher up or to the right or both. Values may be plotted above and below the horizontal axis, and to the left and the right of the vertical axis. An example of such a graph is shown in Chapter 10 (Figure 10.13b, a velocity–time graph for a ball thrown vertically upwards, on page 117).

4.4 Ranges and scales

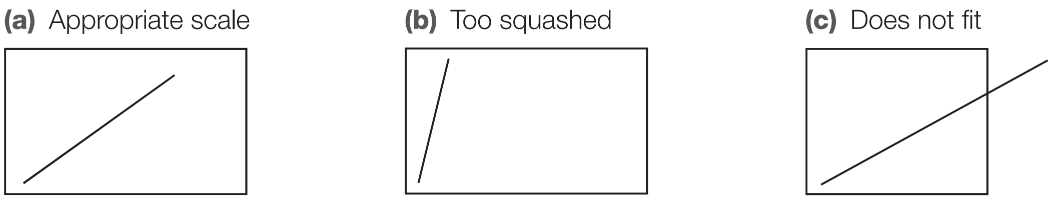


The ***scale*** of an axis is how much each square on the graph paper represents. On the graph paper shown, each main division has 10 small squares. A scale must be chosen for each of the axes so that the ***range*** fits well on the graph paper.

It is helpful if pupils first draw the axes on the graph paper, to ensure that there is sufficient space on the edges of the graph to put labels and values on the axes. They can then see how much space is available for representing the data. (Note that, for simplicity, the edges to the left of and below the axes have not been indicated on the following diagrams.)

Figure 4.4 shows the effect of choosing different scales for the horizontal axis. In Figure 4.4a the scale is appropriate but in Figure 4.4b the data are too squashed and in Figure 4.4c the data do not fit completely on the graph paper.

Figure 4.4 Fitting the range to the graph paper



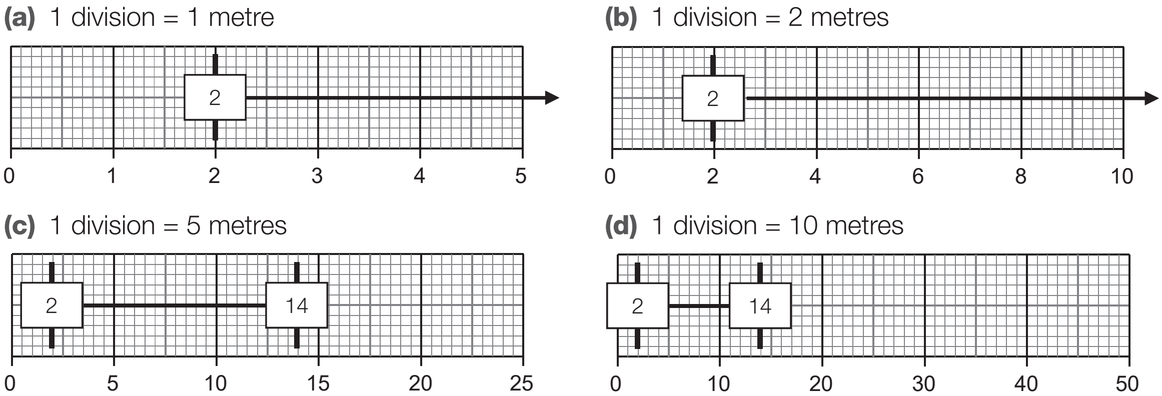
4.5 Choosing a good scale

Pupils find it hard to choose appropriate ***scales***. As well as fitting the range to the graph paper, they need to avoid scales that make the values hard to read. A simple rule is that each large square (main division) should have a value of 1, 2 or 5 multiplied by some power of ten. Other values may be suitable, but this rule works well whether the main squares on the graph paper are divided into 5 or 10 sub-divisions. For this rule, each main square should have one of these values:

	0.1	1	10	100	
etc.	0.2	2	20	200	etc.
	0.5	5	50	500	

This makes it easier to work out the values of the small squares. For example, suppose the ***range*** of values to be plotted on the horizontal axis is from 2 metres to 14 metres. Figure 4.5 shows some possible scales following this rule. Here, the scale in Figure 4.5c would be the best choice.

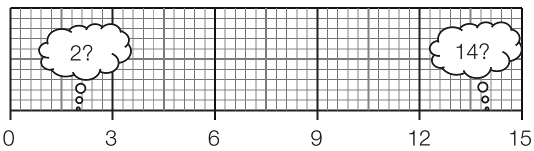
Figure 4.5 Different scales for plotting the range 2 metres to 14 metres



In general, the range of the data on each axis should be *over* a half of the space available. A graph with the data points squashed together makes it hard to read accurate values when interpolating or finding a gradient. You can think of the choice of values for each square (e.g. 1, 2, 5, 10, 20, 50) as a ladder where each step is about twice the previous one. If the range of the data occupies less than half the space on an axis, go up the ladder until it does; if it does not fit on the graph paper, go down a step.

Some pupils try to make the data points fill up as much of the space as possible, by choosing values for the scale divisions that are not on this ladder. This is not a good idea; the scale in Figure 4.6 shows an example (1 division = 3 metres). Although the range of values fits well, it is hard to work out what each small square is worth, so plotting the values is not easy and therefore more likely to lead to mistakes.

Figure 4.6 Another scale: 1 division = 3 metres



On a bar chart, the bars are equally spaced along the horizontal axis, and labelled with the values of the independent variable. Sometimes pupils do something similar on a graph, where the values they put on the horizontal scale reflect the data values so that the data points are equally spaced out. It needs to be emphasised to them that on a graph each scale division has the same value. The convention that the values on each axis increase going up and to the right may also need emphasising. Some pupils may try plotting an axis in the opposite direction if the order of the values in a data table seems to them to be in the 'wrong direction'.

4.6 Labels and units

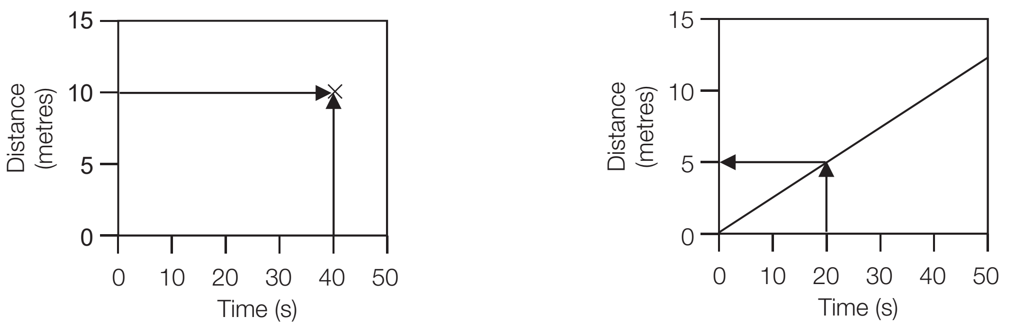
Each axis should include a label that shows the name of the variable and its ***unit***, for example 'Time (min)' and 'Temperature (°C)'. The usual convention in secondary school science is for the units to be enclosed in brackets. It is not recommended that the scientific convention of using the '/' symbol (e.g. 'Time/min') be used until post-16 work. (For more details, see Section 3.1 *Using tables to collect and present data* on page 23.)

4.7 Plotting points and finding values

Being able to use the scales on axes is important for two purposes:

- *Plotting points on a graph*: This involves reading a value on each axis and plotting the point where they cross, as in Figure 4.7a. Each ***data point*** has an ***x-coordinate*** and ***y-coordinate***: the ***coordinates*** determine the position of the data point in relation to the axes.
- *Reading a value off a line*: Once a line graph has been drawn, it can be used to find values at any point along the line (see Section 7.5 *Interpolation and extrapolation on a line graph* on page 70). This involves finding the value on one axis, seeing where it crosses the line, and then reading the value off the other axis, as in Figure 4.7b.

Figure 4.7 Plotting points and finding values



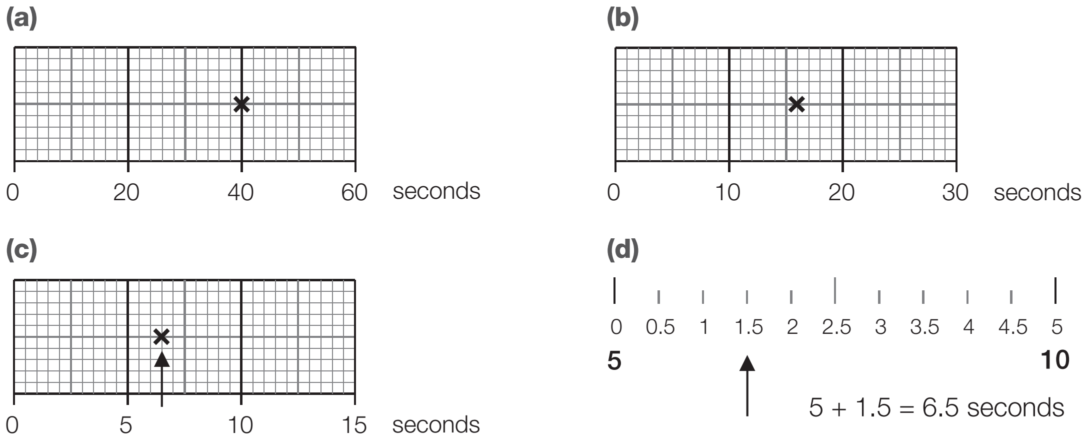
There are no universally accepted conventions for the symbols used to represent data points on graphs. The most common symbol used in school science is ×, but ⊙ and + are also used. To plot a data point, its position can first be marked with a small dot and then diagonal lines drawn through it to form ×, or a circle drawn round it to give ⊙. (Note that sometimes drawing a circle around a data point is intended to indicate an outlier.) The advantage of the + symbol is that the position of the point can be found in two stages by drawing first one line and then the other. For older pupils this can lead in to the idea of drawing error bars to indicate uncertainty in the measurements. The disadvantage is that + may not stand out so clearly visually from the gridlines on the graph paper as ×.

Which symbol (or symbols) to use is really a choice for teachers that depends on the context and what is most appropriate to the pupils. It is a matter of convention not correctness. There is no justification for penalising a pupil for drawing a graph that has 'incorrect' symbols.

4.8 Reading scales

When a data point is on a main division, it is easy to read the value. For example, in Figure 4.8a, the value is 40 seconds. If the data point is not on a main division, it depends on the value of the small squares. In Figure 4.8b, each small square is 1 second, so it is not too difficult to read the value as 16 seconds.

Figure 4.8 Some scales are easier to read than others



The scale in Figure 4.8c is more difficult, since each small square is 0.5 seconds. It may be helpful to jot down the values of the small squares to see how much to add on to the main division (Figure 4.8d). Note that the values for these small squares start at '0'; a common misconception for pupils is that counting always starts at '1'.