# Appendix IV
# How to Create Figures and Tables

Biologists spend their professional lives collecting biological observations which can be presented as numerical facts called ***data*** (singular = ***datum***). Scientific publications are invariably supplied with graphic representations of data. Those arranged in columns and rows of text are known as **tables.** Illustrations consisting of photographs, line drawings or graphs of mathematical functions are known as **figures.**

Each of these types of illustrations is presented, labeled and named following a fixed format. Tables and figures are numbered separately: if there are three tables and three figures in a report, they are numbered Tables 1, 2 and 3 and Figures 1, 2 and 3, respectively.

Every scientific journal has its own requirements for figures and tables, and *UM Publications* (that's us) is no exception. We put forth for you here some guidelines for preparing figures and tables for your own lab reports. Learn these lessons well, and you will have a sound foundation for publishing your own data in the future.

## I. Tables

If your data consist of a series of numbers and/or names and could best be summarized in a series of labeled columns, you should create a **table** to list them. Each column of a table should be clearly labeled, not only with the type of data included, but also with the units of measure, where appropriate.

For example, if you have measured the average tail length of five species of lizard, and wish to list the data in a table, it might look something like Table A4-1. The legend for a table should go above the table itself, as shown.

**Table A4-1. Average snout to vent length (SVL) and tail length of five species of lizards collected in Miami, Florida from June 2000 through June 2001.**

| species | SVL (mm) | tail length (mm) | sample size |
|---|---|---|---|
| *Anolis carolinensis* | 87.1 | 98.2 | 126 |
| *Anolis equestris* | 126.4 | 165.5 | 75 |
| *Anolis sagrei* | 75.1 | 91.2 | 212 |
| *Basiliscus basiliscus* | 185.2 | 207.3 | 87 |
| *Iguana iguana* | 212.1 | 237.4 | 27 |

Notice that the initial column is alphabetical. You need not organize your table in exactly the same way, but it is wise to follow some sort of organization.

A table must always be accompanied by a descriptive legend, located at the *top* of the table, which allows the table to stand on its own. Notice that the legend of Table A4-1 offers more information than that plainly visible in the table. Although you *must* refer to the table in your text, the reader should not have to refer to the text in order to understand the information presented in the table.

In several lab chapters you are provided with tables in which to list your "raw" data. Because these data can be presented as either a table or a graph, you need not do both. The purpose of a scientific report is to elucidate research. It should *not* consist of many complicated-looking pages of redundant data presentation.

# II. Figures

A photograph, line drawing or graph appearing in a scientific paper is called a **figure,** and should be labeled as such.

Most of the figures you will prepare for lab reports in this course will be graphs showing rate functions, or rate curves. For rate curves and any other type of graph, the following guidelines always should be followed.

1. The purpose of a figure is to allow your readers to more easily comprehend your data. Label your axes clearly with the appropriate units of measure. Use the entire page--not just a corner.
2. Each figure must be numbered and be accompanied by a descriptive legend placed *underneath* the figure. Although the figure must be cited in the text of your report, the reader should not have to refer to the text in order to understand how to read your figure. The legend should explain it completely.

The horizontal axis of a graph is known as the **abscissa** or **x-axis**. It is labeled with the units of the **independent variable**. The independent variable is so named because although it changes over the course of the experiment, it is not affected by changes in the experiment. Time is a commonly used independent variable. Its units may be seconds, minutes, hours, months, years, etc.

The vertical axis is known as the **ordinate** or **y-axis**. It is labeled with the units of the **dependent variable**, which changes depending upon the progression of the independent variable. An example of a dependent variable is the change in oxygen volume generated by a chemical reaction over time (independent variable).

Although there are many different types of graphs and graphic functions, the data you will collect in this course will usually be best presented as one of two types of graphs: a **rate curve** or a **histogram**.

## A. Rate curves

Although biological functions may take any number of shapes, you may find that your data often form linear relationships. In order to adequately analyze your results, you must calculate the **slope** of these lines. A slope, with units of y over x, is an expression of **rate**. A rate is an expression of change over an independent variable such as time. "Miles per hour," "millimeters per second," and "pizzas per semester" are all expressions of rate.

For example, if you measured the volume of oxygen gas generated during a chemical reaction (e.g., the breakdown of hydrogen peroxide by the enzyme catalase) for one minute at intervals of 10 seconds, you might obtain "raw" data something like that shown in Table A4-2. ("Raw" data are those straight from the measuring apparatus, which have not undergone any type of mathematical transformation.)

Each of the data points consists of an independent (time) and a dependent (cc $O_2$) variable. These correspond to specific **coordinates**, corresponding to x and y on a graph. The coordinates of the data in Table A4-2, from top to bottom, are (0,0), (10,5), (20,12), (30,20), (40, 37) (50,44) and (60,59). These data are plotted in Figure A4-1.

**Table A4-2. The cumulative volume of oxygen generated by the hydrolysis of hydrogen peroxide by catalase under control conditions (pH 7.0, 25$^{o}$C)**

| time (seconds) | cumulative $O_2$ volume (cc) |
|---|---|
| zero | 0 |
| 10 | 5 |
| 20 | 12 |
| 30 | 20 |
| 40 | 37 |
| 50 | 44 |
| 60 | 50 |

Once you have plotted the data points, study their relationship carefully. Do they form a straight line? An "S" curve? A "J" curve? A parabola? Analyzing data such as these is most accurately done by a computer, but even if you are doing it by hand and eye, there's more to it than "connect the dots." Notice whether the reaction started slowly, picked up speed, then leveled off. If this is the case, the accuracy of your rate calculation will be reduced if you include the more horizontal portions of such an "S" shaped curve. If you are attempting to analyze a reaction which is expected to have a linear rate, you should attempt to draw a straight line through the points which best approximate a straight line.

Notice that in Figure A4-1 the straightest part of the function does not pass through (0,0). Evidently, this particular experimental run started slowly, then increased to a more consistent rate. The dotted line in the figure shows this apparent "J" shaped relationship at the beginning of the experiment. The "best fit" line in Figure A4-1 does not necessarily pass through every data point. The slope of this line equals the rate of the reaction.

Finally notice that the legends of Table A4-2 and Figure A4-1 say exactly the same thing. To include both representations of the same data in your report would be redundant, so choose the best way. In this case, a figure provides more information, and is the better choice.

## Calculating the slope of a line

To calculate the slope of a line, one must determine the change in y (Dy) and divide it by the corresponding change in x (Dx). Because y results in a vertical change and x results in a horizontal change, you may recall that the calculation of slope is sometimes referred to as the calculation of "rise" over "run.". Your rate will be expressed as the units of y over the units of x.

$$\textbf{rate} = \textbf{slope} = \frac{\Delta y}{\Delta x}$$

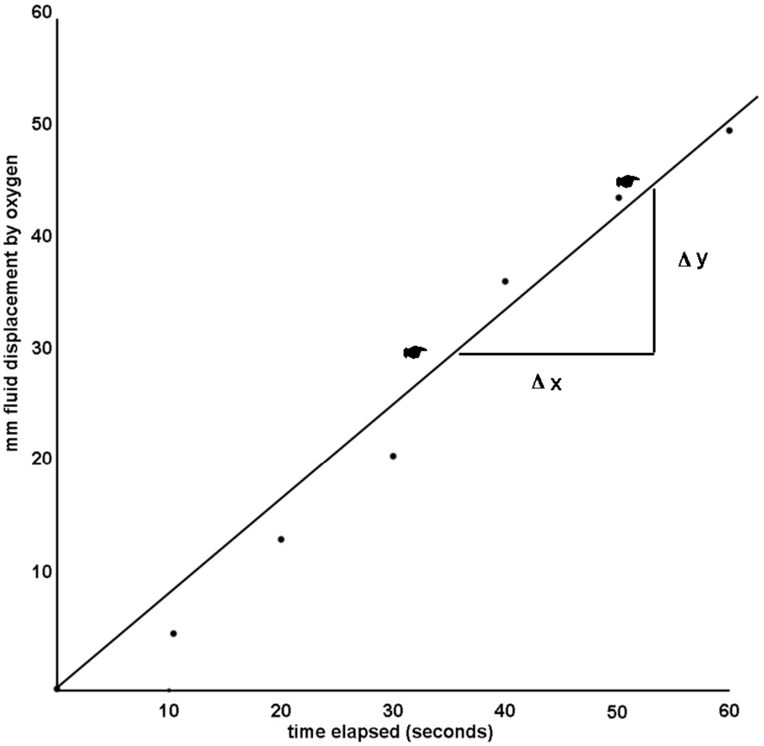

Figure A4-1. Cumulative volume of oxygen generated by the hydrolysis of hydrogen peroxide by catalase under control conditions (pH 7.0, 25°C).

Note that the legend for Figure A4-1 is exactly the same as that for Table A4-1, because the data are the same. In a report, you should publish either the table or the figure, but not both. In this case, a figure is more informative.

In our example, the distance of meniscus rise (proportional to $O_2$ generated and fluid displaced) is plotted against time. Thus, the units of the rate are expressed in mm $O_2$ (y axis) per second (x axis), or more simply, mm $O_2$/sec.

Because the slope of a line is the same no matter where it is measured, you may choose any two corresponding values of x and y. Refer to Figure A4-1 and follow this example to calculate the slope of any rate curve you generate in lab.

1. A "best fit" line through the data points has already been drawn. Notice that the line passes near (but not necessarily *through*) the points which appear to be most linear with respect to each other. Although the first data point should occur at (0,0), the best fit line does not pass through it, apparently because the reaction did not begin immediately at its most consistent rate (i.e., it took a moment to really get going).
2. Choose any two points along the line and determine their coordinates. In Figure A4-1, we have arbitrarily chosen two points with coordinates (35, 30) and (50, 44).
3. Subtract the smaller y value from the larger y value. This quantity is $\Delta y$, or "rise." In our example, $\Delta y = 44 - 30 = \underline{14}$.

4. Next, subtract the smaller x value from the larger x value. This quantity is Δx, or "run." In our example, Δx = 50 - 35 = 15.
5. Divide rise by run (Δy/Δx), being certain to include the units of each variable. The result is the slope of the line, which is equal to the rate of the reaction. In our example, slope = 14 mm $O_2$/15 seconds = 0.9 mm $O_2$/sec.

If you are comparing the rates of several reactions, you may plot them on the same graph, but be sure to differentiate among them by using different symbols or colors. Make your figures easy to read!

If you perform several similar reactions while varying a single physical factor to determine whether that factor affects reaction rate, you should calculate reaction rate (slope) for each experimental run. Next, plot reaction rate as the dependent variable (y-axis) against the independent variable (e.g., temperature, pH, chemical concentration). The resulting figure should show a relationship between reaction rate and the variable. Discuss any trends you observe in this graph in the DISCUSSION section of your lab report.

## B. Histograms

Some types of data are best presented as a **histogram**, or **bar graph**. Histograms are usually used to present frequency data, such as the number of persons in different age classes or the number of frogs in a series of discrete plots along a river. One also may present data in a histogram if it has been collected in discrete time segments (such as 10 minute intervals) during which the important measurement is total change over each time interval.

The *volume* of each bar on a histogram is the important measurement. For example, if you wished to determine the effect of sodium intake on water consumption in mice you might design an experiment in which one group of gerbils is given sodium and another group a placebo. Over the course of an hour, you would measure water consumed every ten minutes. The quantity of water consumed over each ten minute interval would be recorded and later presented in a histogram like the one shown in Figure A4-2.

Notice that at the end of each interval, you should record, not the cumulative water consumption, but only the consumption over that particular time interval. When you analyze the data, add the individual water consumption volumes over each time interval to determine TOTAL water consumption. For example, the total water consumption by control gerbils over the course of one hour is equal to the volume (in cc) of all individual bars together: 1.5 + 0.5 + 0.75 + 1.0 + 1.5 + 1.0 = 6.25. (You may have the opportunity to present data such as this in your own work during the semester. Check the preceding added quantities with those represented in Figure A4-2 to be sure you understand the difference between *interval* water consumption and *cumulative*, or total water consumption.)

A histogram designed this way tells the reader not only the total volume of water consumed by each group in the course of the experiment, but the variation in water consumption between time intervals. Such differences, evident from your graphic data, will allow you to explain differences in water consumption based on differences in behavior and activity, which you also monitored over each time interval.

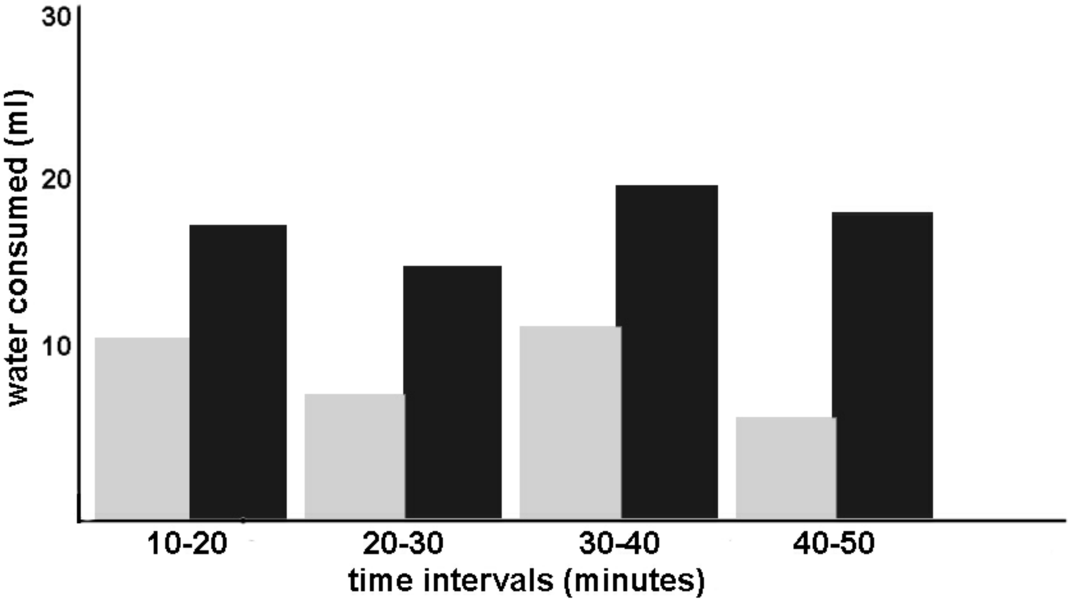

**Figure A4-2. Volume of water consumption by two groups of mice over a series of 10-minute intervals. Black bars indicate gerbils treated with a sodium food supplement and grey bars indicate the control group.**

Note that each bar includes only the water volume consumed over a given 10-minute interval--not the cumulative water consumption up to that point. Do not make the mistake of representing data in a histogram the same way you would on a line graph. They are not the same!

Armed with the information in this appendix and the others, you should have no problem producing a quality laboratory paper on any topic you're assigned. Good luck!