LABORATORY REPORTS: THEIR CAUSE AND CURE

1. Doing the Experiment

Typically your experiment will consist of -

- an idea for how to get the results

- getting some apparatus together (this part is up to you)

- deciding how to set it up to do the measurements (ditto)

- getting it all to work smoothly (This may take the largest fraction of the time you will spend on the project.)

- doing final run-throughs to get your data that you will actually use.

- doing the calculations to go from your *data* to the *conclusions*.

Somewhere in this process, you will probably run into snags. Part of the challenge is to make your particular project work (and no two projects are alike, so everybody's problems will be different).

Now: should you set everything up and then do *all* your measurements before going on to the analysis and theory stage (thinking that this is the best way to stay unbiased by theoretical expectations?) A nice hope, but NO. Before you invest a *huge* amount of your time in the data collection stage, **do a sample run of the complete experiment** first. That is, set up the apparatus, run it, do some sample calculations, and go all the way to the end (including comparison with theory, if that's part of it). You might find that there is something Horribly Wrong with your setup. Or you might find that you don't understand some part of the sequence, even though you thought you did before you started. This "trial run" stage will let you fix things before you go too far.

Getting a good experiment to run from scratch is tough. It takes patience and repetition. Give yourself the time you need to go back and fix things if necessary.

2. What Should Your Writeup Include?

Every experiment that you do in this course will take its final form as a summary report. Like the experiment itself, you can organize the report how you wish. However, in any such writeup, you should provide answers to the following **Five Fundamental Questions**:

- What is the *purpose* of the experiment? That is, what did you intend to do?
- *How* did you conduct the experiment? What were the steps in the experiment, and what were your measurements?
- What was your *result*? How did you go from the raw measurements to the desired endpoint(s)?
- How *uncertain* is your result? What are the most important errors and uncertainties in the measurements, and how do they carry through into the final numbers?
- Finally, if there is some background theory to the experiment, how close is your experimental result to the theoretically expected one?

What other things should you include for a complete writeup?

Photos

Include some photos of your experiment in progress (you, the apparatus, the setup – whatever will show what you actually did).

Data Sheets

Normally, while you are running the experiment you will be writing measurements down as they come. *Include these "raw data" sheets as an appendix to your report.* Don't write down the raw measurements on random scraps of paper and then try to make sense of them later. Organize them right from the beginning in some sensible, tabular order that is easy to sort out and use later for plotting graphs and making calculations.

Graphs

Measurements and data are often presented in the form of graphs. Whether or not you use them is completely up to you – if appropriate to the experiment, then go ahead. There are two kinds:

Histograms: In many cases we are measuring the same quantity over and over again – e.g., the oscillation period T of a pendulum – and we simply want to know how many times we get a value in a certain range. A histogram of the data in this case would be a bar graph of the number of measurements in a certain range of T, plotted versus T. From it, we can deduce a variety of useful results such as the *average* value, the *range* of values, or the *uncertainty* in the average.

Correlation graphs (also called "scatter plots"): In many other cases we have measured two different things that we think are related – e.g., the period T of a pendulum and the length L of its string. Then we can plot a graph of one quantity against the other one (e.g., T versus L) to see if our expected correlation holds up. In this kind of graph we will want to do things like drawing a mean line that seems to fit the data points "best" in some statistical sense. This best-fit line will have a slope and intercept (with uncertainties) that can be read off the graph.

3. The Writing

Scientific writing usually has the purpose of *explaining* or *describing* something – an experiment or derivation. But it doesn't need to be stereotypically dry and dull. Compare the following two descriptions of the same thing:

"We released the ball from a measured height at the top of the ramp. We used a stopwatch to measure the time for the ball to roll to the floor, taking an average of six trials all released from the same height."

"The ball was placed at the top of the ramp, measuring its height above the floor carefully with a meter stick on each trial. The ball was then released at the same moment as the stopwatch was started, and the stopwatch was allowed to run until the moment when the ball hit the floor. The total time which was taken by the ball to reach the bottom of the ramp was then recorded. Six trials were performed, and comparing the individual values of each trial, it was determined that the average of the six gave a reliable result for the elapsed time."

Which one would you rather read if you were marking a report?

In the second one, the description is accurate and understandable – that is, it gives a correct outline of what actually happened. That's good. But it is convoluted, long, and ugly. That's bad. Notice that it is all in *passive voice*: "The ball was placed", "It was determined", and so on. The people doing the experiment have disappeared into limbo. By contrast, the first version is a lot shorter, just as understandable, and far more appealing to read. Notice that it is written in *active voice:* "we released the ball", etc. It is perfectly all right to bring in the personal pronouns "we" and "I" where appropriate. You're real people, and so is your reader.

The bottom line is to write clear, concise, and accurate descriptions of what you did. Stick to the topic and keep it simple, direct, and active.

3. Measurement Uncertainty:

How good are your measurements? You will need to say something about that in your writeup.

No measurement of a physical quantity is exact or perfect. Every measurement is **uncertain** to some degree and, in most experiments, knowing *how uncertain* the result is is just as important as the result itself. So, you will need some way to evaluate this measurement uncertainty during your experiment.

There is no totally universal rule to use here, but very often the best way is to let the measurements themselves tell you how uncertain they are. You can *repeat your measurements* a bunch of times and see just *how* repeatable they are. There are actually two good reasons for doing this:

- A few repeated measurement will usually show any "wild" values that went wrong for any number of reasons, and give us more security about the right range of the quantity,
- Averaging several measurements will give us a more *precise* answer than just one measurement; i.e. it will give us a result with lower uncertainty.

How do we actually do this?

Let's take the simple case where you have a repeated series of measurements $\{x_1, x_2, x_3, ...x_n\}$. The **average** or **mean** is

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

Now, $\overline{x}$ is a better estimate of the "true" value of the quantity we are trying to get than any single measurement would be.

But how *much* better? That is, what is the uncertainty of $\overline{x}$?

The answer comes from basic statistics theory, and is called the **standard deviation** of the measurements $\{x_i\}$. We will denote this new quantity as s_x: to calculate it, roughly speaking we take the difference between each x_i and the mean $\overline{x}$, average the squares of all those differences, and take the root of the result. The actual formula defining it is

$$s_x = \Big(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2\Big)^{1/2}$$

It is also called the *rms [root-mean-square] scatter* of the measurements (for reasons which become obvious when you look at the formula). *The standard deviation s_x represents the typical amount by which the individual measurements differ from the average.*

The formula for s_x makes it look a bit tedious to calculate for a long list of numbers. However, most calculators (including the "McMaster standard" calculator) make the job easy. Any calculator which has built-in statistical functions will allow you to enter a series of numbers, and then press one or two specially defined keys which immediately give you $\overline{x}$ and s_x. Learn how to do this with your calculator or within a spreadsheet package.

The one other bit of help we need from statistical theory is that *the average of n measurements is better than a single measurement by the square root of the number of measurements.* So, the **uncertainty of the mean** is

$$\sigma_{\overline{x}} = \frac{s_x}{\sqrt{n}}$$

Thus, if you want to know $\overline{x}$ with half the uncertainty of a single measurement, you need to make *four* independent measurements; if you want to make the uncertainty 5 times smaller, you need 25 measurements; and so on.

In summary: do several repeat measurements. Then, their standard deviation s_x will give you a good estimate of the uncertainty in any one measurement. And, the uncertainty of the mean $\sigma_{\overline{x}}$ will tell you how precise the final average is. That's what you need to know.

Error: Accuracy vs. Precision

We now need to discuss a different phenomenon: *error*.

An effect that might occasionally creep into the experiment, and which is often much harder to evaluate properly, is called **systematic error** or **bias**. It implies that something has actually gone wrong in the experiment, and does *not* mean the same thing as as the *random* process of measurement uncertainty.

For example, think about timing measurements with a stopwatch: even though the stopwatch itself is perfectly accurate, the person using it may tend (for example) to start it always just a little too late because of slow reaction time, or just a little too early because of too-eager anticipation. If this happens consistently over all the timing trials, all of your time measurements may be systematically wrong by small amounts, in addition to the random differences.

Systematic error is a bad thing if it lurks undetected. To prevent it, or find out if it is present, you could (e.g.) try to repeat the experiment in a different way, or with different equipment, or with different people (for example, to get around the stopwatch timing problems mentioned above, you might get two or three people to make the timings and see how different the results are).

Understand your own experiment, and make it work.

Two different terms – "precision" and "accuracy" – are used to refer to "measurement uncertainty" and "error". They are different things. A measurement is called **precise** if it has *low random uncertainty*. On the other hand, a measurement is called **accurate** if it has *negligible systematic error*, that is, if the "true" value is within its range of uncertainty.

Your report should show that you have thought about (and tried to avoid) errors and biasses. However, don't go into far-fetched discussions with long lists of unrealistic possible sources of error. And do not –*ever* – use the term "human error". It's a cover-up meaningless term.