A little book about motion

Lesson 2: Being certain about your uncertainties

> "The principle of science, the definition, almost, is the following: The test of all knowledge is the experiment. The experiment is the sole judge of scientific 'truth."
>
> — Feynman, *The Feynman Lectures on Physics*, Vol I, 1-1

Figure 1: Legendary physicist Richard Feynman is seen here at the Nobel Prize Award ceremony in 1965. He won the Nobel Prize in physics for his contributions to the development of quantum electrodynamics and he was a keen populariser of physics through books and lectures.

All measurements are uncertain to some degree and drawing scientific conclusions depends very much on how certain you are about your uncertainties. Most students (and people in general) find it very difficult to wrap their heads around uncertainty. There is no easy "formula" to follow when estimating and dealing with uncertainties and every experiment is associated with its own unique uncertainties. We have general approaches and sensible rules to follow, but in the end, students have to actually *think critically and independently* about what they are doing. The real world is messy and confusing[1], and knowing how to cope with that is a very important skill. Modern science was invented – the human race truly grew up – when we realised this.

Let's perform a simple motion experiment and use it to go over a few basic concepts: Throw a tennis ball up in the air. Try to throw it up with the same amount of force so it reaches the same height (this is difficult to control so this "experiment" is not very well-designed – we'll discuss that later). Have another person measure the time it takes from the moment it leaves your hand to the moment you catch it again. Repeat the experiment 5 times. These repetitions are called **trials**. Open up a spreadsheet and type your values in different columns. Figure 2 shows my trials after I typed them in. The **header** contains the *name* of the variable I'm measuring and the *unit* of measurement: 'Time / s'. I've merged five cells, so the header is common to all trials. The unit s (for seconds) applies to all trials.[2] When considering the uncertainties of this

[1] "Confusion is the sweat of learning" - it's good if you are confused! It means you have started the process of learning.

[2] The slanted line in 'Time / s' is actually a division sign and the logic is as follows: Time is a physical quantity expressed as a value and a unit, e.g. "Time = 0.82 s", so if we only want the numerical value we divide both sides by the unit and get "Time / s = 0.82".

A	B	C	D	E
Time / s				
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
0.9	0.82	0.87	0.83	0.94

Figure 2: The raw data.

experiment, we first need to consider the uncertainty due to the equipment used. Take a look at your stopwatch or timer that you used. What is the smallest unit that the timer can measure? On a typical stopwatch, it is 0.01 s (or one centisecond, $1 \times 10^{-2}\,\mathrm{s} = 1\,\mathrm{cs}$, although nobody ever says that – still, you ought to know the prefixes listed in table 1) and we could call this the **equipment uncertainty**. One sensible rule would be to say that *the equipment uncertainty on a digital scale is equal to this smallest unit*, so e.g. if the measurement was 0.90 s, then we will assume that this measurement is likely to be any value between 0.89 or 0.91 seconds. We write

nano	n	10^{-9}
micro	μ	10^{-6}
milli	m	10^{-3}
centi	c	10^{-2}
kilo	k	10^{3}
mega	M	10^{6}
giga	G	10^{9}

Table 1: Some of the common **prefixes** that we use in science to modify the size of a unit.

$$0.90 \pm 0.01\,\mathrm{m}$$

When a measurement is expressed in this way, we call 0.90 the **best estimate** and 0.01 the **absolute uncertainty**. Notice how we write 0.90 (not just 0.9) since we are measuring with an uncertainty on the second decimal so we need to show two decimals in our recording of the best estimate. Hence 0.90 is written with two **significant digits**, whereas 0.9 only has one significant digit[3]. Since every measurement was taken with this stopwatch, we add this equipment uncertainty to our table as shown below (you get the ± symbol on a mac by pressing 'option' + 'shift' + '+'). I also right-aligned the cells so that the decimals line up nicely and I increased the number of decimal places in the first cell so that it shows 0.90, not just 0.9 (your spreadsheet has a button for doing this - see if you can find it!).

[3] In science, zeroes to the left of a number don't count as significant (who would ever write 37 as 0037?), but zeroes to the right of a number do. If there is any doubt, use the scientific notation $a \times 10^{b}$ to clarify things.

A	B	C	D	E
Time / s				
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
± 0.01	± 0.01	± 0.01	± 0.01	± 0.01
0.90	0.82	0.87	0.83	0.94

Figure 3: The raw data with equipment uncertainty and showing the correct number of significant digits.

Now consider the first measurement:

$$0.90 \pm 0.01\,\mathrm{m}$$

How "good" is this measurement? Well, "good" depends on many factors, but one thing we can do is quantify how **precise** this measurement is: Take the absolute uncertainty 0.01 and divide it by the best estimate 0.90 to get the **relative uncertainty**. If you multiply this by 100%, you get the corresponding **percentage uncertainty**:

$$\frac{0.01}{0.90} \approx 0.011 \approx 1\%$$

We often *round off percentage uncertainties to the nearest integer* so they are easier to read. What does 1% mean? It means that the uncertainty of what I'm measuring is only 1% of the measured quantity. Any measurement with a percentage uncertainty less than 1% would be considered very precise. An uncertainty between 1% and 5% could be considered precise, between 5% and 10% a bit imprecise, more than 10% rather imprecise, and so on. Those words, however, are subjective and therefore meaningless. *What matters is that you have quantified the uncertainty so that people can make up their own minds about it*. Would you get on an airplane if you knew 1% of them crashed?[4] What about 0.001%? Would you happily take a vaccine if 1% (1 out of 100) of those who took it died from a blood clot? What about if 2% (2 out of 100) of people died from the disease itself? When your life depends on it, you want results that are extremely precise!

[4] The actual probability of a plane crashing is around 0.00002%. Source: This link.

Precision is one thing, but **accuracy** is something else. Consider our time measurements – do they actually represent the time it takes the ball to go up and down? Maybe you stopped the stopwatch too early so that the actual time is a bit longer? It might be a very precise measurement (1%), but if you are not measuring the variable accurately, then all that precision isn't particularly useful! If you stopped the stopwatch too early on every trial, then we call that a **systematic error**: All the measured time values are shorter than what the actual time is. Systematic errors cause all measurements to be inaccurate in the same way (e.g. if you forget to zero an electronic balance when measuring mass, then all mass measurements will be off by the same amount). We always strive for our scientific measurements to be *precise and accurate*. Figure 4 is an illustration of the difference between precision and accuracy.

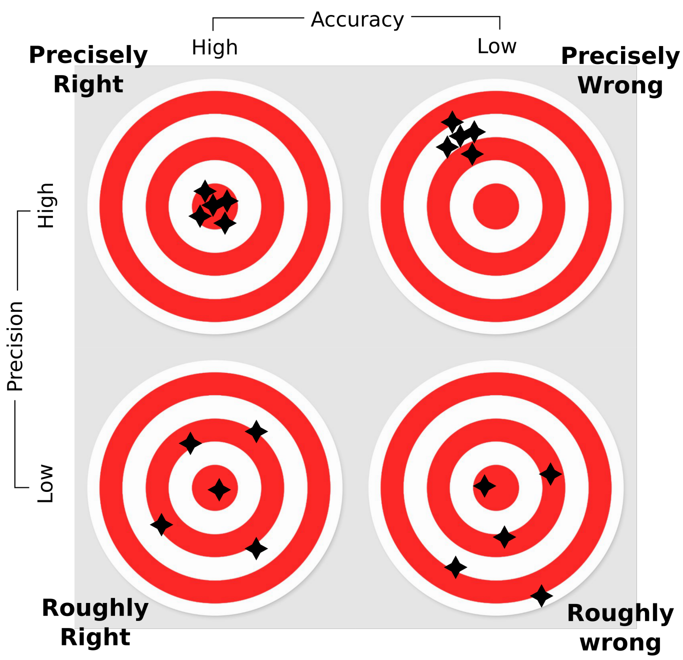


Figure 4: Comparing the two different concepts of *precision* and *accuracy*.

Imagine you measure the density[5] of a cube to check if it's aluminium. In your first experiment, you measure

$$2.64 \pm 0.01\,\mathrm{g/cm^3}$$

In your second experiment, using a different method, you measure

$$2.63 \pm 0.18\,\mathrm{g/cm^3}$$

You can visualise the range of these two measurements using **error bars** as shown in figure 5. Say you know that the density of pure aluminium is $2.70\,\mathrm{g/cm^3}$ by looking up the value in a well-respected physics book (we call that the **table value** or **accepted value**). Which experiment is most precise? Which one is accurate? Which experiment would you say is "best"? Discuss!

[5] **Density**, denoted by the Greek letter 'rho' ρ, is defined as mass per unit volume:

$$\rho \equiv \frac{m}{V}$$

The unit is typically given in $\mathrm{g/cm^3}$ or $\mathrm{kg/m^3}$. Table 2 lists a few densities.

Substance	Density / $(\mathrm{g/cm^3})$
water (at 0°C)	0.9998
water (at 5°C)	1.0000
rock (granite)	2.7
iron	7.87
mercury	13.5
osmium	22.6
atomic nuclei	2.3×10^{14}
a black hole	2×10^{27}

Table 2: Densities of some substances. One of the interesting properties of water is that the density depends on temperature. Osmium is the densest *element*, but by far the densest objects in the the Universe are atomic nuclei, neutron stars, and black holes.

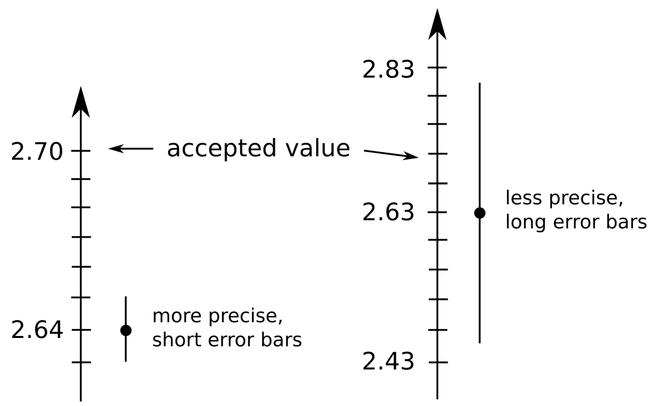


Figure 5: Error bars are used to visualise the uncertainty of a measurement.

I was once asked to measure the density of a cube of aluminium. I measured it with different methods and got successively more precise measurements, but at some point I noticed that the table value of the density started to lie outside my range of error (like the short error bars shown above). The more precise my measurements became, the more inaccurate the result. My teacher was also puzzled by it, but in the end we discovered that the cube we thought was pure aluminium was actually a metal alloy containing a small amount of another metal! So in the end my measurements turned out to be precise and accurate *and* they revealed something new about the world – *which is the way science works*.

Currently (2021) there is a fascinating "Crisis in Cosmology" (also dubbed "Hubble Trouble"): Cosmologists measure the expansion rate of the Universe by measuring something called Hubble's Constant. Different teams use different methods and as those measurements get more and more precise, they deviate from each other more and more! The heart of the matter is that 67.4 ± 0.5 and 73.2 ± 1.3 disagree (the ranges don't overlap). Are these measurements precise? Which measurement is inaccurate? What is going on? Nobody currently knows! We do know, however, that we don't fully understand this aspect of the Universe and hence we need to collect more data and use that to modify our theories about how the Universe works.

Let's return to our tennis ball experiment. We have measured the time it takes for the ball to go up and down and we have performed five trials. We now need to condense those trials down to just one measurement and we also need to find the uncertainty of that. The easiest way to do this is to calculate the arithmetic mean (the **average**) of the trials and the **standard deviation** of the trials. In most spreadsheets you can easily calculate both by using the "`average`" and "`stdev`" formulas. You go to the cell to the right of all your trials, write "`=AVERAGE(`", and select the cells of your trials. Likewise with "`=STDEV(`":

fx =STDEV(A4:E4)

A	B	C	D	E	F	G
Time / s						
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Average	Std. dev.
± 0.01	± 0.01	± 0.01	± 0.01	± 0.01		0.04970 ×
0.90	0.82	0.87	0.83	0.94	0.872	=STDEV(A4:E4)

Figure 6: Use the built-in spreadsheet functions to easily calculate the average and standard deviation of the trials.

Above, you can see the standard deviation is initially given with lots of significant digits (0.04970 . . .). *We always round absolute uncertainties off in a sensible way and that typically means to one significant digit*[6], in this case that would be 0.05 (your spreadsheet will have a button that decreases the number of decimal places - see if you can find it). *Since* 0.05 *is rounded to the second decimal, then the average needs to be rounded off to the second decimal place too.*[7] So in our example, the average would be rounded off to 0.87 s with an absolute uncertainty of 0.05 s:

$$0.87 \pm 0.05\,\mathrm{s}$$

The table ends up looking like the one below (notice how I merged the cells in the top row, so that "Time / s" is again a header for all relevant columns):

[6] The exception to this rule (there are always exceptions, don't get too attached to certain rules!) is that if the first significant digit is a 1, then you are allowed to include a second significant digit. E.g. 73.2 ± 1.3 is acceptable, as is 2.63 ± 0.18.

[7] Notice how *the number of significant digits is often NOT equal to the number of decimal places*. If I got one dollar every time a student made that mistake, I would be a millionaire :)

A	B	C	D	E	F	G
Time / s						
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Average	Std. dev.
± 0.01	± 0.01	± 0.01	± 0.01	± 0.01		
0.90	0.82	0.87	0.83	0.94	0.87	0.05

Figure 7: The final processed data.

When we calculate the percentage uncertainty of the average quantity, it is less precise than just one single measurement:

$$\frac{0.05}{0.87} \approx 0.057 \approx 6\%$$

Didn't we just say that the uncertainty of the equipment was 1%? Why is the uncertainty of the average six times bigger? Well, the equipment uncertainty is still 1%, but *other, more unpredictable, uncertainties are revealed when we take trials*. In our experiment we can point to a few rather obvious explanations for these uncertainties:

1. The most serious issue is that when designing this "experiment" we assumed we could throw the ball up in the air with the same force so it went to the same height. This assumption[8] is not satisfied since it's very difficult to always throw a ball with exactly the same force and that will surely introduce some more uncertainty. For example, a ball thrown up with a larger force will have a longer time of flight. (This is why I said it was a poorly designed experiment in the first place.)

2. Another issue is that it's hard to tell exactly when I let go of the ball and when I catch it because the motion happens so quickly. Since these are the moments when I start and stop the stopwatch, we would expect this lack of certainty to also introduce some random uncertainty. Also, I could be catching it at different points in its trajectory, sometimes a little bit higher, sometimes a little bit lower, etc.

[8] In scientific experiments (and life in general!) we always make a number of assumptions (consciously or unconsciously) that are worth questioning from time to time.

There are probably more issues/weaknesses to discuss, some more significant than others, but the main point to learn here is that there will always be **random uncertainties** that are revealed by taking trials – *the whole purpose of taking trials is to reveal these random uncertainties*.

When the random uncertainty is larger than the equipment uncertainty (as in our case above, 6% > 1%), then it's important to discuss how to reduce those random uncertainties. If, on the other hand, the random uncertainty is smaller than the equipment uncertainty (see below), then the random uncertainties are negligible and the precision of your data is only limited by the precision of your equipment. In any case, it's important to remember that *the largest uncertainty always "wins", small uncertainties "drown in the big ones"*.

After identifying the causes of the random uncertainties, you always need to suggest ways of improving those issues/weaknesses. Here are ways in which we could improve our experiment in order to reduce the mentioned random uncertainties:

1. *Instead of throwing the ball up by hand*, we should use some sort of projectile launcher that has a fixed mechanism for launching a ball with the same force every time. At school, we actually do have such a projectile launcher that has a built-in spring that compresses a certain distance every time and will therefore launch a ball with the same force every time.

2. *Instead of relying on our eyes (and moving hand) to detect when the ball is released and caught*, we could set up a light gate that can detect when the ball passes through it on the way up and on its way down (alternatively, we could film the motion and use video software to identify the exact frames at which the ball is launched and when it comes down again).

Let's see how these improvements would affect the results: Below is data that I collected by using a projectile launcher and a light gate that measured the time of flight more consistently. The time values are different from the tennis ball experiment because the launcher shoots out a little metal sphere that doesn't go up very high. The light gate time measurements had an equipment uncertainty of 0.002 s:

Time / s						
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Average	Uncertainty
± 0.002	± 0.002	± 0.002	± 0.002	± 0.002		0.0014
0.541	0.541	0.541	0.538	0.539	0.540	=STDEV(E4:I4)

Figure 8: After improving the experiment we get much more precise data. Since the standard deviation of the trials is smaller than the equipment uncertainty, we have eliminated all significant random uncertainties.

Notice that the standard deviation of these five trials is actually *less than* the equipment uncertainty ($0.0014 < 0.002$). Hence, the equipment uncertainty overrules the random uncertainty (the largest uncertainty always "wins"), and we can confidently state that *we have eliminated all the random uncertainties for this particular experiment*. In order to further improve the precision (which is in fact not necessary because now it's very precise, see (1) below), we would need to use better equipment. Since it doesn't make sense to have an uncertainty of the average that is less than the precision of the equipment used, I will in this case replace the standard deviation with the equipment uncertainty (0.002 s):

Time / s						
Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Average	Uncertainty
± 0.002	± 0.002	± 0.002	± 0.002	± 0.002		
0.541	0.541	0.541	0.538	0.539	0.540	0.002

Figure 9: The final table with raw and processed data.

With these two improvements, the percentage uncertainty of the average decreased to

$$\frac{0.002}{0.540} \approx 0.004 \approx 0.4\% \tag{1}$$

which is much lower than 6% (more than 10 times – one *order of magnitude* – smaller). Clearly the precision improved with these changes.

Lesson 2: Questions and activities

1. **Spreadsheets.** Go to the following data tables, make a copy, and use the spreadsheet functions to calculate (i) the average, (ii) the uncertainty of the average, (iii) and the percentage uncertainty of the average. Also, don't forget to round off all values appropriately and fix any mistakes you find!

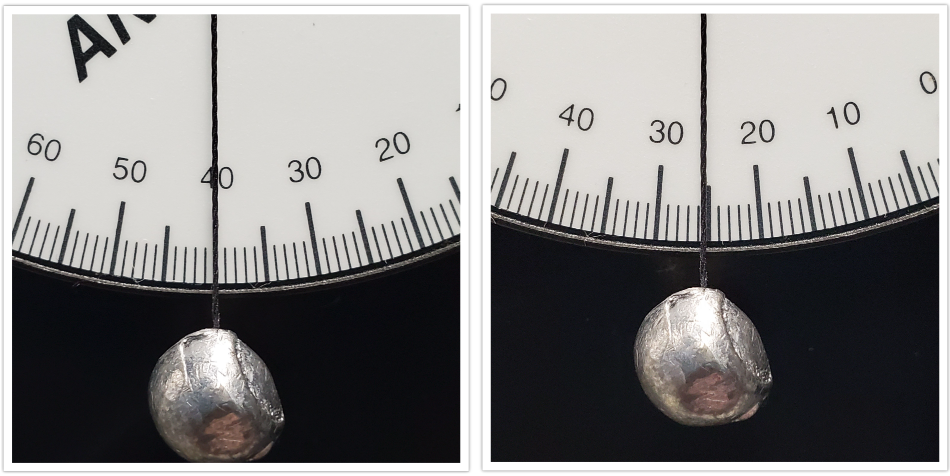
(a) Spreadsheet 1.

(b) Spreadsheet 2.

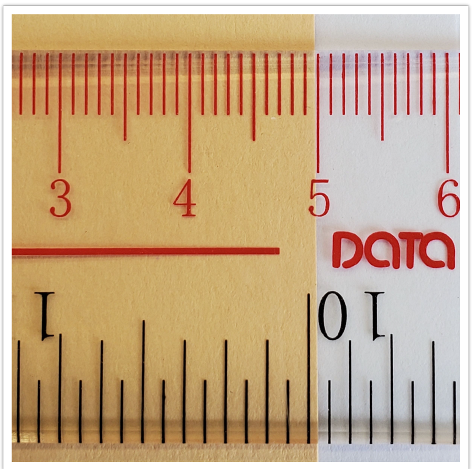
2. The density of many types of rock is around $2.7\,\mathrm{g/cm^3}$ but the average density of planet Earth is around $5.5\,\mathrm{g/cm^3}$. What conclusion can you draw from this?

3. **Analog scales.** It was mentioned in the text that the uncertainty of a *digital* scale is typically just the smallest unit displayed. For *analog* devices, it's a bit more tricky. Here we need to take the spacing of the lines into account too. The usual rule is to take *half the smallest unit as the absolute uncertainty* but it really depends on the spacing on the scale! Here are a few pictures of analog scales. Determine the best estimate of the measurement, the absolute uncertainty, and the percentage uncertainty:

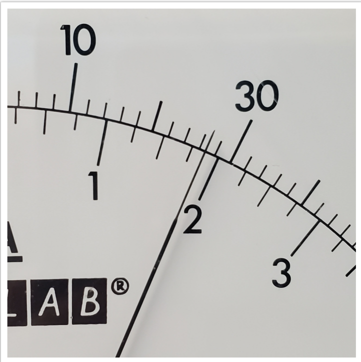
   (a) A protractor measuring in degrees:

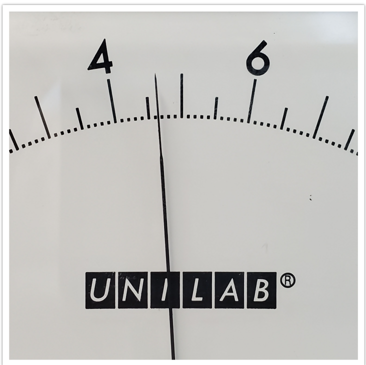


   (b) Finding the edge of some yellow paper using a ruler that measures in millimeters ($1\,\mathrm{mm} = 1 \times 10^{-3}\,\mathrm{m}$):



(c) Current measurements using analog ammeters that measure in milliamperes (mA). The first picture has two scales because the equipment can be set to operate at two different scales. Estimate against both scales:





4. **Length:** When measuring the length of an object, you often need to measure two positions and then find the difference between those two position measurements. Since both measurements are uncertain *the overall absolute uncertainty of the length measurement is the sum of the two absolute uncertainties* (consider the minimum and maximum possible lengths and this makes sense). This is one example of 'error propagation': When performing calculations with measured quantities, the uncertainty 'propagates' (= travels) through the calculations. So if the two position measurements are

$$x_1 \pm \delta x_1 \quad \text{and} \quad x_2 \pm \delta x_2,$$

then the length will be

$$L = x_2 - x_1$$

with an absolute uncertainty

$$\delta L = \delta x_1 + \delta x_2$$

For example, if

$$x_1 = 0.0 \pm 0.5\,\text{mm} \quad \text{and} \quad x_2 = 7.5 \pm 0.5\,\text{mm}$$

then

$$L = 7.5 \pm 1.0\,\text{mm}$$

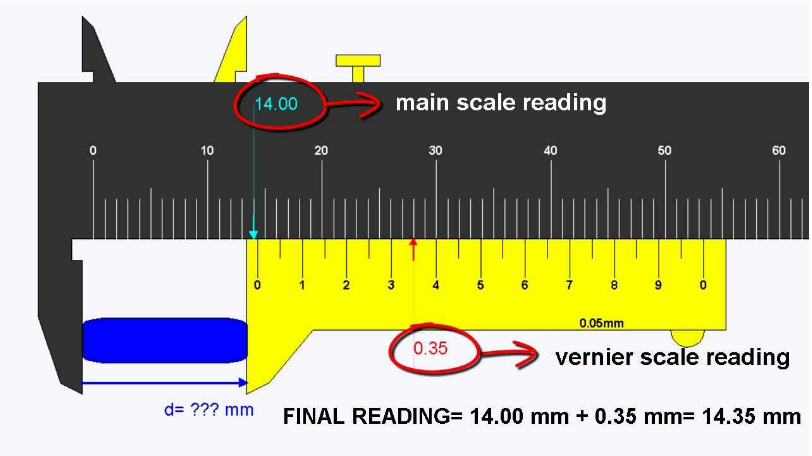
Perform the following measurements, estimate the equipment uncertainty and calculate the percentage uncertainty of each one:

(a) With a normal ruler: The width of a sheet of A4 paper. (Rulers and paper are behind the whiteboard)

(b) With a normal ruler: The width of your pen

(c) With a normal ruler: The width of your laptop charger wire

(d) Explain why the above three measurements get more and more imprecise.



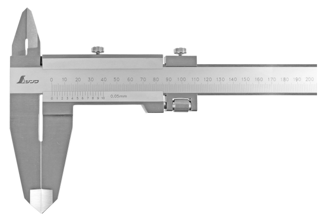


Figure 10: A vernier calliper allows you to measure well-defined lengths more precisely.

(e) With a vernier caliper: The width of your pen

(f) With a vernier caliper: The width of your laptop charger wire

(g) How many times more precise is the vernier caliper compared to a normal ruler?

(h) **Length** is a dimension that is usually measured in the fundamental (base) SI unit **meter**. Look up the current definition of what a meter is.

5. **Mass:** Use a digital scale to measure the mass of the following items. Again, estimate the equipment uncertainty and calculate the percentage uncertainty:

(a) A sheet of A4 paper.

(b) Five randomly chosen 50 g masses from a 0.5 kg mass hanger set. Work out the average and percentage uncertainty of the average.

(c) Take out two large glass beakers. Put the first one on the scale and zero it. Use the second beaker to pour water into the first and use the scale on the beaker to measure off 0.5 L = 500 mL of water. How sure are you that this is an accurate value of the mass of 500 mL of water?

(d) $1\,\mathrm{L}$ is per definition $1\,\mathrm{dm}^3$. How many liters is one m^3? And how many liters is one cm^3?

(e) Measure the mass of ONE drop of water (use a plastic dropper). The density of water is $1\,\mathrm{g/cm}^3$. What is the volume of ONE drop of water in mL?

(f) Measure the mass of 50 DROPS of water and use that to calculate the mass of ONE drop of water - what is the volume of ONE drop of water according to this measurement? Is this a more precise measurement than (e)?

(g) **Mass** is a dimension that is usually measured in the fundamental (base) SI unit **kilogram**. Look up the current definition of what a kilogram is (it changed recently!).

6. **Time:**

   (a) Go to this website and measure your reaction time. Estimate the uncertainty of each measurement, perform 5 trials in a spreadsheet, work out the average and the uncertainty of the average.

   (b) Measure how long it takes to count off 30 seconds in your head. Take 5 trials, work out the average, uncertainty, and percentage uncertainty. Are you systematically counting too fast or too slow?

   (c) **Time** is a dimension that is usually measured in the fundamental (base) SI unit **seconds**. Look up the current definition of what a second is.

7. **Liquid adhesion:** Take the following measuring cylinders from the cupboard: 10/0.2 mL, 25/0.5 mL, and 100/1 mL. Pour some water in them and measure the volumes as precisely as possible (measure to the bottom of the convex meniscus). Estimate the equipment uncertainty and calculate the percentage uncertainty of your measurement.

8. **The seven base SI units:** There are seven fundamental (base) SI units, three of which were mentioned above (meter, kilogram, and second). Explore this link, list the other four base SI units and write up their definitions.

   Answers to all the questions.

Lesson 2 Quiz

Check your understanding of this lesson: Here is a quiz.