Errors and uncertainties in biology internal assessment

Biological systems are complex and difficult to control. Nevertheless, biological investigations require measurements to be made, and biology students need to be aware of the sources of error in their data, both qualitative and quantitative. For the purposes of internal assessment, work assessed for data collection and processing must contain quantitative data suitable for processing. The expectations with respect to errors and uncertainties in internal assessment are the same for both standard level and higher level students, and are supportive of topic 1.1 of the subject guide.

The treatment of errors and uncertainties is directly relevant in the internal assessment of:

- data collection and processing, aspects 1 and 3 (recording raw data and presenting processed data)
- conclusion and evaluation, aspects 1 and 2 (concluding and evaluating procedures).

Expectations at standard level and higher level

An appreciation of errors should be apparent at all stages of a report on an investigation:

- in the **design** stage, where the limitations of time and the materials should be assessed, and the potential sources of error should be controlled. The magnitude and significance of normal (background) variation in biological systems should be appreciated.

- in the **data collection and processing** stage, where the degree of accuracy of a measuring device should be stated as well as other observed sources of error

- in the **conclusion and evaluation** stage, where the sources of error should be discussed, along with possible ways of avoiding them.

- Although students should analyse their investigations for sources of error, they should not be led to conclude that, with all such sources of error and imprecision, experimental results are worthless. Experimental results are only **estimates**.

Terms and concepts in error analysis

(a) Random variation or normal variation

In biological investigations, errors can be caused by changes in the material used, or by changes in the conditions under which the experiment is carried out. Biological materials are notably variable. For example, the water potential of potato tissue may be calculated by soaking pieces of tissue in a range of concentrations of sucrose solutions. However, the pieces of tissue will vary in their water potential, especially if they have been taken from different potatoes. Pieces of tissue taken from the same potato will also show variations in water potential, but they will probably show a normal variation that is less than that from samples taken from different potatoes. Random errors can, therefore, be kept to a minimum by careful selection of material and by careful control of variables. For example, you could use a water bath to reduce the random fluctuations in ambient temperature.

Human errors can become random when people have to make a large number of tedious measurements and, therefore, their concentration spans vary. Automated measuring, using a data logger system, can help to reduce the likelihood of this type of error. Alternatively, the experimenter can take a break occasionally.

(b) Human errors (mistakes)

Human errors can occur when tools, instruments or protocols are used or read incorrectly. For example, a temperature reading from a thermometer in a liquid should be taken after stirring the liquid and with the bulb of the thermometer still in the liquid. Thermometers (and other instruments) should be read with the eye level with the liquid in the thermometer (reading needle) to prevent parallax error. Human errors can be **systematic**, because the experimenter does not know how to use the apparatus properly, or they can be **random**, because the power of concentration of the experimenter is fading.

(c) The act of measuring

When a measurement is taken, this can affect the environment of the experiment. For example, when a cold thermometer is put into a test tube with only a small volume of warm water in it, the water will be cooled by the presence of the thermometer, or when the behaviour of animals is being recorded, the presence of the experimenter may influence the animals' behaviour.

(d) Systematic errors

Systematic errors can be reduced if equipment is regularly checked or calibrated to ensure that it is functioning correctly. For example, a thermometer should be placed in an electronic water bath to check that the thermostat of the water bath is correctly adjusted. A blank should be used to calibrate a colorimeter to compensate for the drift of the instrument.

Accuracy
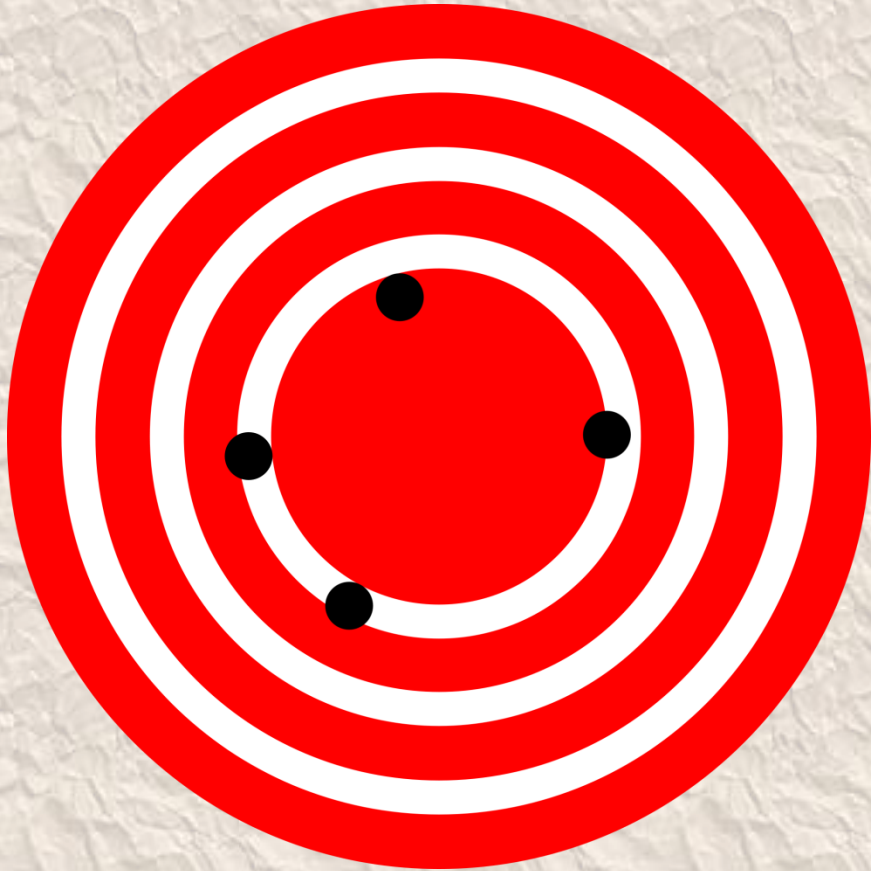
Accuracy describes the closeness of arrows to the bull's-eye at the target centre. Arrows that strike closer to the bull's-eye are considered more accurate. The closer a system's measurements are to the accepted value, the more accurate the system is considered to be.

High accuracy, but low precision

Precision

When all arrows are grouped tightly together, the cluster is considered precise since they all struck close to the same spot, even if not necessarily near the bull's-eye. The measurements are precise, though not necessarily accurate. However, it is *not* possible to reliably achieve accuracy in individual measurements without precision—if the arrows are not grouped close to one another, they cannot all be close to the bull's-eye. Their *average* position might be an accurate estimation of the bullseye, but the individual arrows are inaccurate.)

High precision, but low accuracy

(e) Degrees of precision and uncertainty in data

Students must choose an appropriate instrument for measuring such things as length, volume, pH and light intensity. This does not mean that every piece of equipment needs to be justified, and it can be appreciated that, in a normal science laboratory, the most appropriate instrument may not be available.

For the degrees of precision, the simplest rule is that the degree of precision is plus or minus (±) the smallest division on the instrument (the **least count**). This is true for rulers and instruments with digital displays.

The **instrument limit of error** is usually no greater than the least count and is often a fraction of the least count value. For example, a burette or a mercury thermometer is often read to half of the least count division. This would mean that a burette value of 34.1 cm^3 becomes 34.10 cm^3 (± 0.05 cm^3). Note that the volume value is now cited to one extra decimal place so as to be consistent with the uncertainty.

The **estimated uncertainty** takes into account the concepts of least count and instrument limit of error, but also, where relevant, higher levels of uncertainty as indicated by an instrument manufacturer, or qualitative considerations such as parallax problems in reading a thermometer scale, reaction time in starting and stopping a timer, or random fluctuation in an electronic balance read-out. Students should do their best to quantify these observations into the estimated uncertainty.

Other protocols exist for recording uncertainties. In biology internal assessment (IA), no specific protocol is preferred, and a moderator will support a teacher when it is clear that recording of uncertainties has been required and the uncertainties are of a sensible and consistent magnitude.

(f) Propagating errors

Propagating errors during data processing is not expected but it is accepted provided the basis of the experimental error is explained.

(g) Replicates and samples

Biological systems, because of their complexity and normal variability, require replicate observations and multiple samples of material. As a rule, the lower limit is five measurements, or a sample size of five. Very small samples run from 5 to 20, small samples run from 20 to 30, and big samples run from 30 upwards. Obviously, this will vary within the limits of the time available for an investigation. Some simple investigations permitting a large sample, or a large number of replicate measurements, could be included in the scheme of work to reinforce this point. It is also possible to use class data to generate sufficient replicates to permit adequate processing of the data. However, each student must have been personally involved in the data collecting process, and their own set of raw data should be presented and clearly identified.

Where sufficient replicates have been carried out, then the calculation of the standard deviation of the mean is expected. Another statistic, the standard error of the mean to derive confidence limits, may also be calculated. The standard error is not expected, but it would be an acceptable alternative to the standard deviation.

In order to establish the significant difference between two samples, it may be possible to calculate a student's t-test. However, this would not be systematic as it is only appropriate to use this statistic when certain conditions apply (interval data, sample sizes greater than five, normal distribution of the population).

Where these statistics are calculated from a preset menu on a calculator or computer, a worked example will not be expected, but the data should be presented in such a way that the steps in the processing can be clearly followed.

Students should be made aware that, if a reading is particularly different from the others, it may be left out of the processing and analysis. However, students must **always** justify why they have chosen to do this.

Interpreting the relevant assessment criteria

Data collection and processing: Aspect 1 (recording raw data)

In tables of raw data, the degrees of precision of a measuring instrument should be given at the head of a column along with the units (see part (e) above).

The number of decimal places in the raw data should agree with this degree of precision.

It may be that, in spite of extensive searching, the student does not have access to the degree of precision of a measurement, for example, a solution prepared by a supply company or an instrument that lacks technical specifications.

Tables 1–4 below show the raw data from an experiment that compared the behaviour of strips of potato and apple tissues all cut to 4 cm long then soaked in different sucrose solutions.

Table 1: DCP aspect 1 = "not at all"

The results

Solution	Potato measurements	Apple measurements
0.2	4/3.8/4.2/4.1/4.1	4.1/4.2/4.2/4.1/4.2
0.4	3.8/3.7/3.7/3.7/3.8	4.1/4.2/4.3/4.2/4.2
0.6	3.8/3.7/3.7/3.8/3.6	4/4/4.1/4.1/4
0.8	3.6/3.5/3.7/3.7/3.5	4.1/4/3.9/3.9/4
1	3.7/3.6/3.7/3.7/3.6	3.8/4/4/3.8/3.9
Distilled water	4.2/4/3.9/4/4.2	4.2/4.3/4.1/4.3/4.4

The data is badly organized, there are no units or uncertainties, and it is too ambiguous to be comprehensible.

Table 2: DCP aspect 1 = "partial"

The results

Sucrose	Potato lengths					Apple lengths				
0	4.2	4	3.9	4	4.2	4.2	4.3	4.1	4.3	4.4
0.2	4	3.8	4.2	4.1	4.1	4.1	4.2	4.2	4.1	4.2
0.4	3.8	3.7	3.7	3.7	3.8	4.1	4.2	4.3	4.2	4.2
0.6	3.8	3.7	3.7	3.8	3.6	4	4	4.1	4.1	4
0.8	3.6	3.5	3.7	3.7	3.5	4.1	4	3.9	3.9	4
1	3.7	3.6	3.7	3.7	3.6	3.8	4	4	3.8	3.9

The table contains appropriate quantitative data. The title is inadequate but the data avoids total ambiguity as it has correct column headings. There are no units or uncertainties given and no associated qualitative data is recorded. The number of decimal places in the data is variable. This is something that programs like MS Excel® do by default unless the student knows how to use the control on the toolbar to set the number of decimal places. So even though a piece of tissue may measure exactly 4 cm, it should still be recorded as 4.0 cm.

Table 3: DCP aspect 1 = "partial"

The lengths of potato and apple tissues after soaking.

Sucrose/ mol dm$^-_3$	Potato lengths/ cm					Apple lengths/ cm				
0	4.2	4.0	3.9	4.0	4.2	4.2	4.3	4.1	4.3	4.4
0.2	4.0	3.8	4.2	4.1	4.1	4.1	4.2	4.2	4.1	4.2
0.4	3.8	3.7	3.7	3.7	3.8	4.1	4.2	4.3	4.2	4.2
0.6	3.8	3.7	3.7	3.8	3.6	4.0	4.0	4.1	4.1	4.0
0.8	3.6	3.5	3.7	3.7	3.5	4.1	4.0	3.9	3.9	4.0
1.0	3.7	3.6	3.7	3.7	3.6	3.8	4.0	4.0	3.8	3.9

The table contains appropriate quantitative data with units. The title is not very precise but would be sufficient. However, there are no uncertainties and no associated qualitative data were recorded.

Table 4: DCP aspect 1 = "complete"

Lengths of two plant tissues, potato (*Solanum*) and apple (*Malus*) after soaking in solutions of sucrose of different concentrations. The initial lengths were 4.0 cm.

Sucrose/ mol dm^{-3}	Potato lengths/ cm ± 0.1 cm					Apple lengths/ cm ± 0.1 cm				
0	4.2	4.0	3.9	4.0	4.2	4.2	4.3	4.1	4.3	4.4
0.2	4	3.8	4.2	4.1	4.1	4.1	4.2	4.2	4.1	4.2
0.4	3.8	3.7	3.7	3.7	3.8	4.1	4.2	4.3	4.2	4.2
0.6	3.8	3.7	3.7	3.8	3.6	4.0	4.0	4.1	4.1	4.0
0.8	3.6	3.5	3.7	3.7	3.5	4.1	4.0	3.9	3.9	4.0
1	3.7	3.6	3.7	3.7	3.6	3.8	4.0	4.0	3.8	3.9

It was also noticed before the soaking that the potato tissue floated in the solution from 0.4 to 1.0 mol. The apple tissue, however, only floated in the solutions from 0.6 to 1.0 mol. After soaking the tissues became softer at higher sucrose concentrations but they were quite hard in the lower concentrations.

The student has designed a table where the appropriate data are organized clearly with units and uncertainties. The table has a precise title and there is relevant associated qualitative data recorded.

Data collection and processing: Aspects 2 and 3 (processing raw data and presenting processed data)

These two aspects will often be assessed on the same table or graph.

Processing data in biology often requires a statistical analysis of the data. This is because of the inherent variability of the material used as well as variation due to its manipulation. Thus the previous set of data (table 1) will possess an uncertainty because of the instrument used to measure it (a millimetre ruler), the dexterity of the experimenter in cutting, and the variability in the potato and apple tissue. A student could represent this by calculating a margin of error. The simplest would be plus or minus the range of measurements or plus or minus half the range of measurements. If the data permits, the error margin could be represented by plus or minus the standard deviation of the mean or the standard error of the mean. These ranges may be expressed as error bars on graphs. Though this is not obligatory, it would support assessment statement 1.1.1.
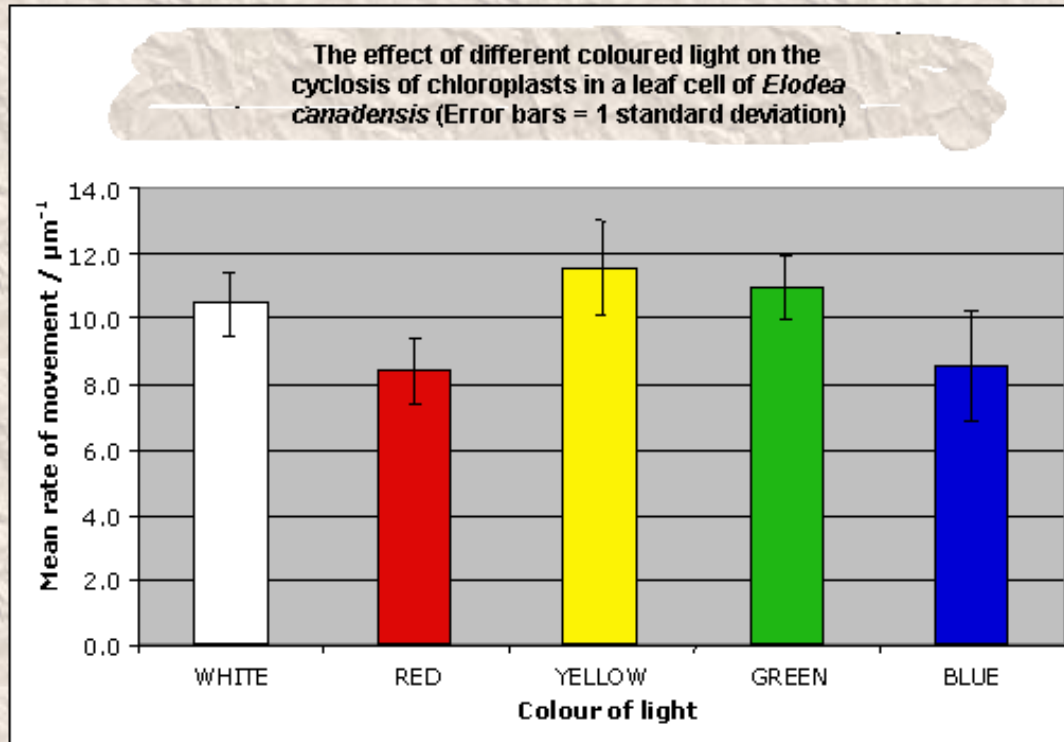
Table 5: DCP aspect 2 = "partial"

*Table 5 is from an investigation studying the effect of the colour of different light on the movement of chloroplasts in Elodea leaf cells by cyclosis. It shows both raw data and processed data. Here the appropriate units and degree of precision are given in the title of the column. The rates of movement of the chloroplasts (this is the raw data in this case, as it is not calculated but measured) are recorded to the appropriate number of decimal places. Unfortunately, the means and standard deviations do not respect these degrees of precision. This student would only achieve **partial** for aspect 2 of data collection and processing*

	Speed of chloroplast / $\mu m\ min^{-1}$				
Trial	White	Red	Yellow	Green	Blue
1	12.6	9.3	14.1	10.6	8.6
2	10.3	6.4	11.6	10.2	11.7
3	9.9	7.5	13.4	12.5	8.2
4	11.1	8.1	12.1	10.8	7.0
5	9.1	9.5	10.4	11.0	9.9
6	10.8	9.8	10.6	12.5	9.7
7	10.0	8.3	10.4	9.7	6.4
8	10.2	8.4	9.9	10.3	8.7
9	10.5	8.4	11.6	11.0	7.0
Means	10.5	8.41	11.56	10.95	8.57
Standard deviations	0.972	1.054	1.436	0.968	1.682

The speed was estimated using a micrometer eyepiece calibrated for high power (×400) on the microscope. At this magnification each graduation represented 2.6 μm.

Figure 1: DCP aspect 2 = "partial"

This processing could be accompanied by a bar chart (figure 1) to present the data more clearly.

The effect of different coloured light on the cyclosis of chloroplasts in a leaf cell of *Elodea canadensis* (Error bars = 1 standard deviation)
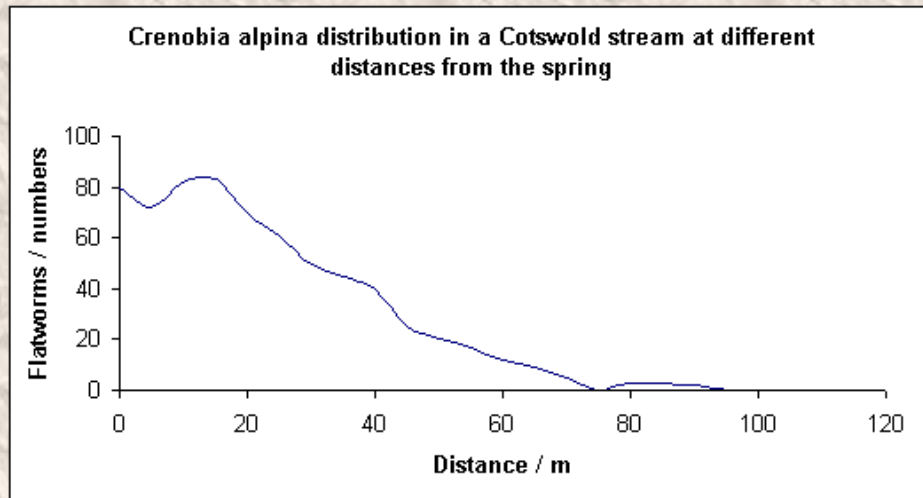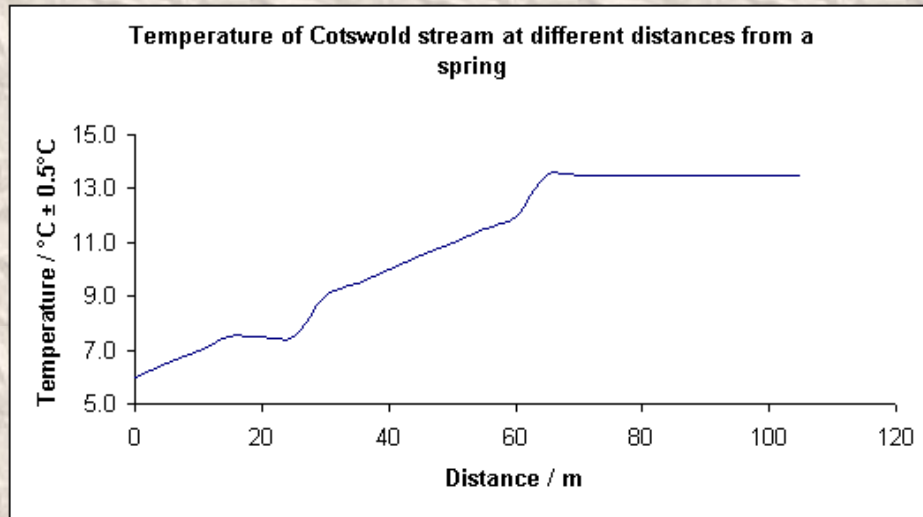
The graph on its own would be awarded **partial** for aspect 2 of data collection and processing, as the steps in the processing should be shown. The tabulated data would be required for a "complete" for aspect 2. However, a worked example of the calculation of the mean or the standard deviation would not be required, as these are part of a calculator statistics menu.

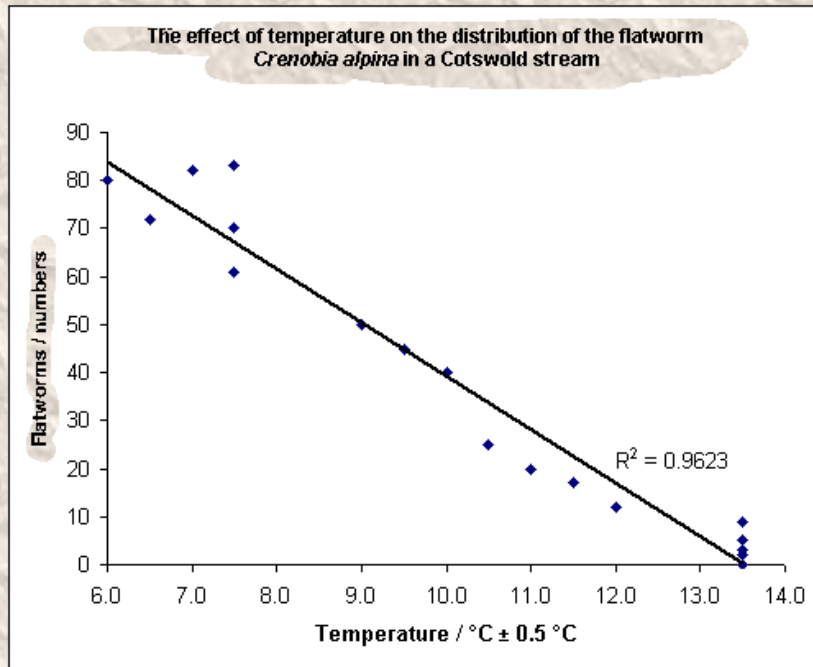# Figures 2 and 3: DCP aspect 2 = "partial", aspect 3 = "complete"

Students who simply reproduce their raw data as graphs (as shown in figures 2 and 3) cannot expect more than partial to be awarded for aspect 2 of data collection and processing. The graphs in figures 2 and 3 show the data obtained from a field study of the distribution of a species of flatworm in a stream.

Figures 2 and 3: DCP aspect 2 = "partial", aspect 3 = "complete"

Temperature of Cotswold stream at different distances from a spring

Crenobia alpina distribution in a Cotswold stream at different distances from the spring

Uncertainties are given where they are relevant, but the raw data has only been plotted on a graph. The graphs are drawn using the correct conventions and they have a clear title. This would permit the student to make some conclusions about the relationship between water temperature and the distribution of the animal, but not with much certainty. The student would score partial for aspect 2 and complete for aspect 3 of data collection and processing.

Figure 4: DCP aspect 2 = "complete", aspect 3 = "complete"

The effect of temperature on the distribution of the flatworm *Crenobia alpina* in a Cotswold stream

$R^2 = 0.9623$

Flatworms / numbers (y-axis)

Temperature / °C ± 0.5 °C (x-axis)

Here the student has fully processed the data and would achieve complete for aspect 2 of data collection and processing. The relationship between the variables is plotted as a scatter plot and a trend line has been drawn. The computer program has also been used to calculate a correlation coefficient, which can be used to back up the conclusion. It would not be expected for the student to comment on the degree of significance of the correlation, just that it is high or low. However, some appreciation that larger samples can have lower coefficients and still be significant would be acceptable. For aspect 3 the correct conventions are used for the graph. It has a clear title and uncertainties are shown through both the use of the trend line and the degree of precision on the x-axis. This would be awarded a complete for aspect 3 of data collection and processing.

The set of graphs in figures 5–8 show a range of responses to the data given in table 4.

Sucrose/ mol dm^{-3}	Potato lengths/ cm ± 0.1 cm					Apple lengths/ cm ± 0.1 cm				
0	4.2	4.0	3.9	4.0	4.2	4.2	4.3	4.1	4.3	4.4
0.2	4	3.8	4.2	4.1	4.1	4.1	4.2	4.2	4.1	4.2
0.4	3.8	3.7	3.7	3.7	3.8	4.1	4.2	4.3	4.2	4.2
0.6	3.8	3.7	3.7	3.8	3.6	4.0	4.0	4.1	4.1	4.0
0.8	3.6	3.5	3.7	3.7	3.5	4.1	4.0	3.9	3.9	4.0
1	3.7	3.6	3.7	3.7	3.6	3.8	4.0	4.0	3.8	3.9

Figure 5: DCP aspect 2 = "partial", aspect 3 = "not at all"
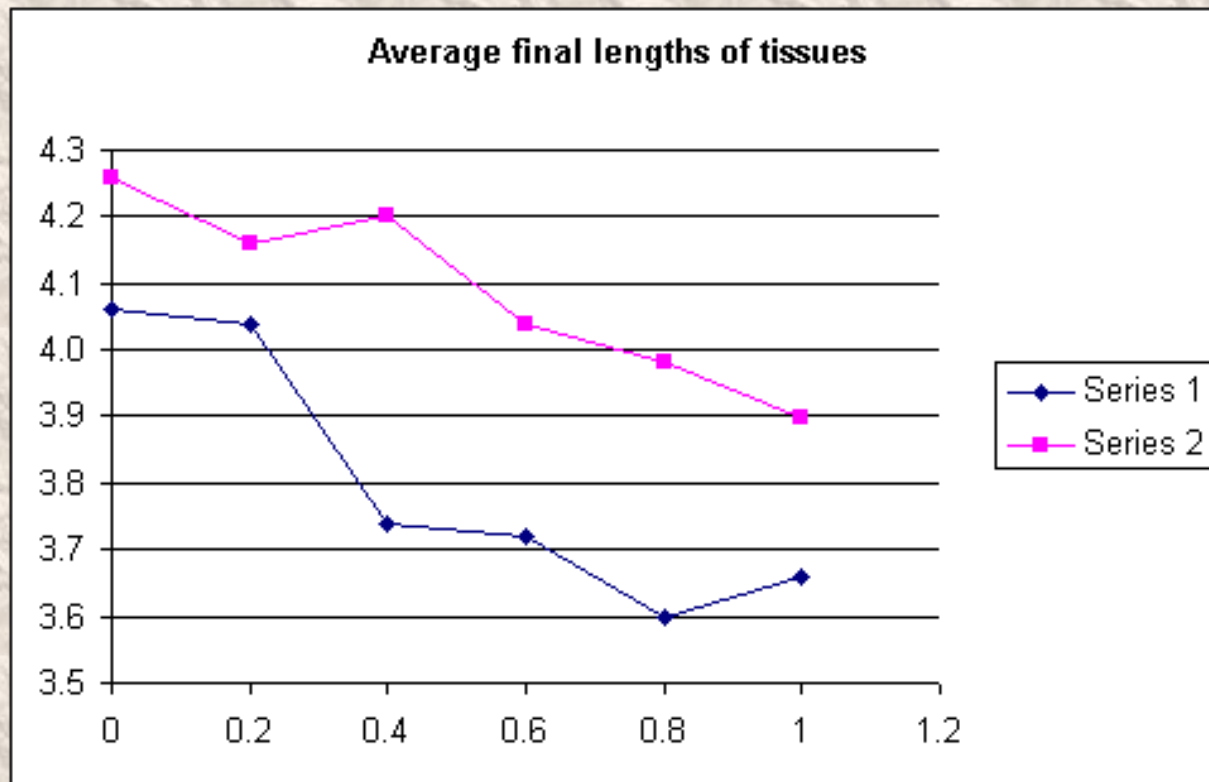
Some processing has been carried out as expressed in the title, although the table shown in the calculations would need to be given. This work can be awarded partial for aspect 2. The graph, however, is incomprehensible. There are no labels on the axes and the key for the lines does not reveal which tissue is which. As a result, not at all is awarded for aspect 3 of data collection and processing.

Figure 6: DCP aspect 2 = "partial", aspect 3 = "partial"

Average final lengths of potato and apple tissues
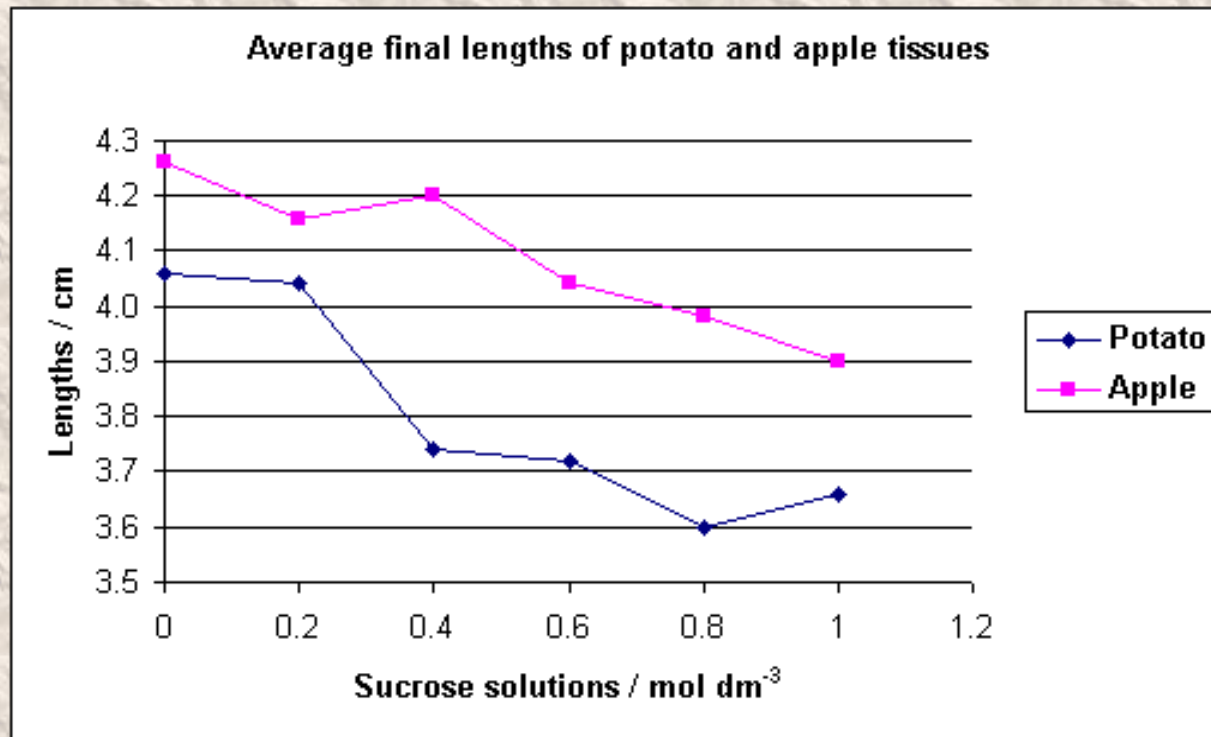
Here there is evidence of some relevant processing (the average lengths), so the work can be awarded partial for aspect 2 (again, a table showing the steps in processing would be required).

The graph is drawn using the correct conventions and it has an adequate title but there are no uncertainties shown. Aspect 3 would be awarded partial.

Figure 7: DCP aspect 2 = "partial", aspect 3 = "partial"

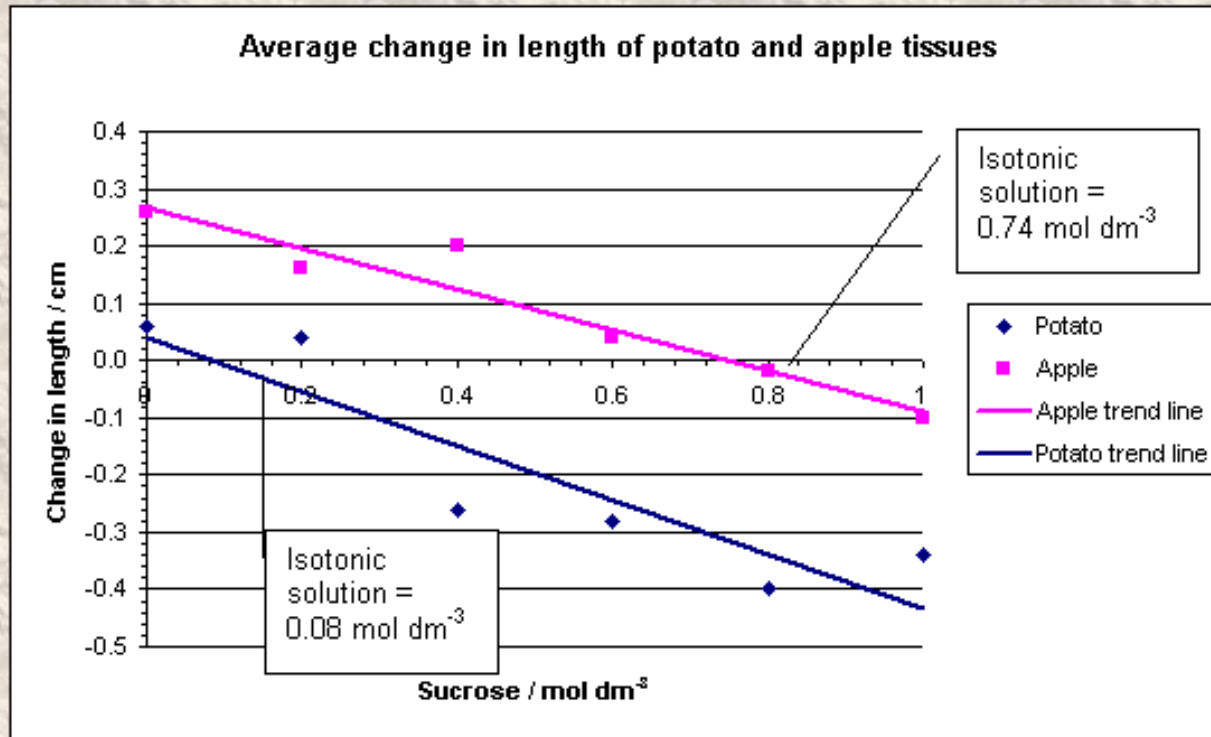Average change in length of potato and apple tissues
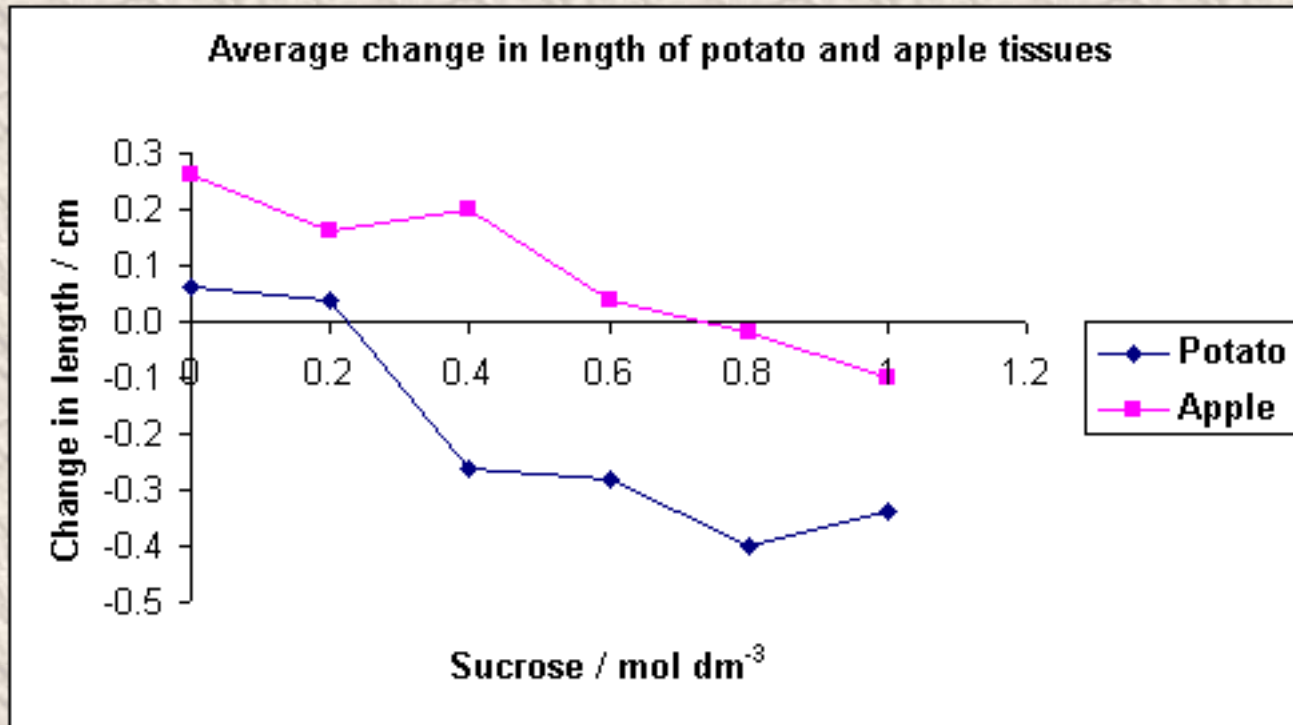
Though most of the processing has been done (calculating the average change in lengths), unfortunately the comparison of the two tissues is not complete without the estimation of the isotonic solutions. This can only be awarded partial for aspect 2.
The graph is drawn using correct conventions and the title is appropriate, but there are no uncertainties given as degrees of precision, as error bars or as a trend line. Therefore only partial can be awarded for aspect 3 of data collection and processing.

Figure 5: DCP aspect 2 = "complete", aspect 3 = "complete"

Average change in length of potato and apple tissues

Isotonic solution = 0.74 mol dm^{-3}

Isotonic solution = 0.08 mol dm^{-3}

- ◆ Potato
- ■ Apple
- — Apple trend line
- — Potato trend line

Change in length / cm

Sucrose / mol dm^{-3}

The graph is used to process the data by using the intercepts of the x-axis to establish the sucrose solutions isotonic to the tissue sap. So long as this graph is accompanied by a table showing how the data was processed, aspect 2 would achieve complete. Note in this method that the calculation of the percentage change in length is not necessary as all the pieces of tissue were cut to the same initial lengths (4.0 cm).

The graph is drawn using the correct conventions, it has a clear title and uncertainties are given by the use of trend lines. The axes have graduations on them to improve the precision of the estimate of the isotonic solutions. Aspect 3 has achieved complete.

Conclusion and evaluation

Errors and uncertainties are often relevant in aspects 1 and 2 of the conclusion and evaluation criterion because students are expected to reach a reasonable and justified interpretation of the data, and to appreciate the quality of the procedure (producing a measure of precision and accuracy).

Aspect 1 (concluding)
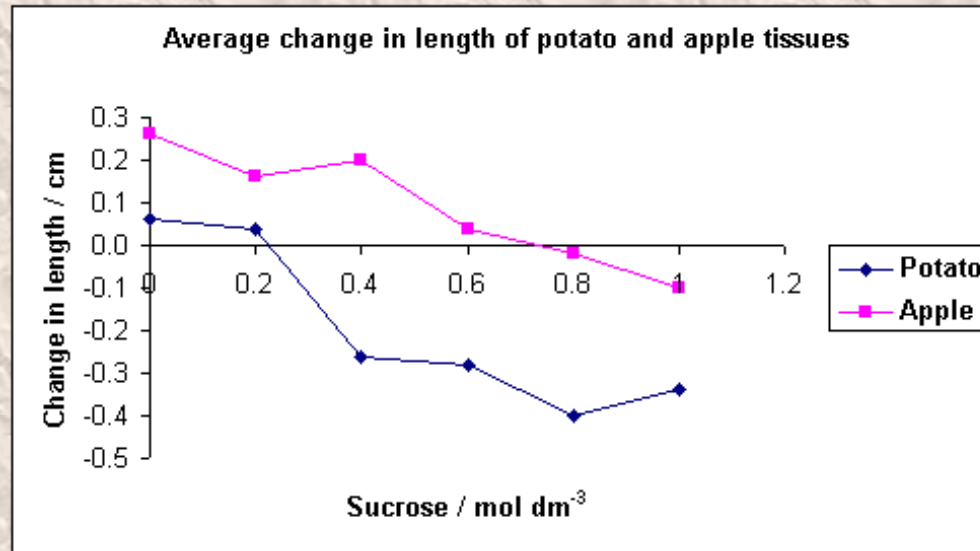
A student who drew figure 7 above under aspect 3 of data collection and processing does not construct a best straight line and makes an approximate estimate of the isotonic solution. Under aspect 1 of conclusion and evaluation, the following earns a not at all.

"The graph shows that the sucrose solution where the potato tissue does not change in size is about 0.2 mol dm^{-3} whereas the solution where the apple tissue does not change in size is about 0.7 mol dm^{-3}. Therefore the apple tissue contains more sucrose than the potato tissue."

Not only does the student ignore any variation in the data, but a serious error is made in the conclusion. Sucrose is not necessarily the only solute in the tissue sap, if it is there at all.

Aspects 2 and 3 (evaluating procedures and improving the investigation)

The evaluation of the investigation must take into account the weaknesses in the method that undermine the conclusion that can be reached. The criticisms must also include weaknesses observed in the data. If these variations in the data can be linked to observations made and noted during the investigation (for example, in a log book), it reveals that the student is truly evaluating the investigation and not desperately looking for something to say.

In the course of an investigation to investigate the diffusion of methylene blue stain through agar gel, a student obtained the results and took the photograph displayed in figure 9, and made the following statements in the evaluation based upon the quantitative and supporting qualitative data.

The diffusion of methylene blue of different concentrations in agar. Error bars = ± 1mm
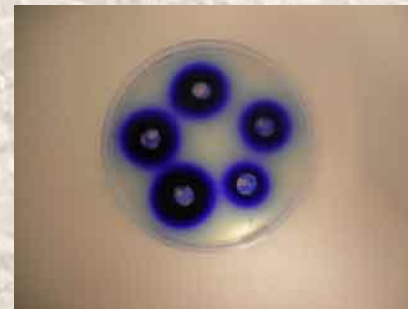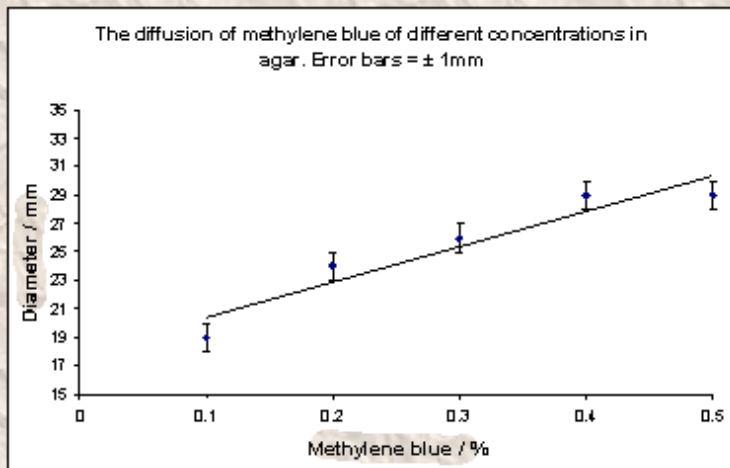
"The variations observed in the data could be due to the zones of different concentrations on methylene blue overlapping and interfering with one another (as can be seen on the photograph). A bigger dish or separate dishes of agar could be used to avoid this happening.

The shape of the diffusion zones is not always circular so the result will depend upon which way the diameter is measured. This could have been due to inconsistencies in the gel, it was observed that small bubbles were trapped in it and when it was cut using the cork borer the cracks developed in the gel round the edges of the hole. A sharper cutting tool and a more uniform gel would be better.

It was also difficult to tell exactly where the blue colour stopped as it faded into the surrounding gel. So the measurements are a bit subjective. Good lighting and a uniform background are recommended when measuring.

The range of concentrations was limited to 0.1–0.5%. A wider range may show whether the straight line relationship is in fact a curve."

This evaluation would be awarded complete for both aspects 2 and 3 of conclusion and evaluation. The criticisms made are based upon the measurements and observations made during the investigation. The suggested improvements are realistic and would probably lead to improved results. Note that these may not be the only criticisms that could be made, but an exhaustive list is not what is expected.