The following text was adapted from the VCAA .

Measurement terms

VCE Chemistry requires that students can distinguish between and apply the terms ‘accuracy’, ‘precision’, ‘repeatability’, ‘reproducibility’ and ‘validity’ when analysing their own and others’ investigation findings. An understanding of the terms ‘accuracy’ and ‘precision’ is also important in the analysis and discussion of investigations of a quantitative nature.

Accuracy

A measurement result is considered to be accurate if it is judged to be close to the ‘true’ value of the quantity being measured. The true value is the value (or range of values) that would be found if the quantity could be measured perfectly. For example, if an experiment is performed and it is determined that a given substance had a mass of 2.70 g, but the true value of mass is 3.20 g, then the measurement is not accurate since it is not close to the true value. The difference between a measured value and the true value is known as the ‘measurement error’.

‘Accuracy’ is not a quantity and therefore cannot be given a numerical value. It is allowable for a measurement to be described as being ‘more accurate’ when its method and/or instruments clearly reduce measurement error, such as using a triggered electronic timer system compared to a hand-operated stopwatch. Accuracy may not be quantified: ‘measurement error’ is the quantity used to evaluate how close a measured value is to the true value.

While accurate measurements and observations are important in all science experiments, in some cases it may not be possible to determine the accuracy of a measurement since a true (or accepted) value for a physical quantity may be unknown at the conditions under which the experiment is conducted. For example, the accepted value of the ionic product of water of 1.0×10^{-14} M^2 only applies at 25 ºC. This value does not apply at other temperatures. As a result, the pH of pure water is 7.0 only at 25 ºC. Many practical activities in the classroom involve an experimental setup that is unique to the student; for example, determination of the conductivity of the water in the water tank or dam on the student’s property. In such instances, there is no accepted single value with which comparisons can be made.

Precision

Experimental precision refers to how closely two or more measurement values agree with each other. A set of precise measurements will have very little spread about their mean value. For example, if a given substance was weighed five times, and a mass of 2.70 g was obtained each time, then the experimental data are precise. However, this gives no indication of how close the results are to the true value and is therefore a separate consideration to accuracy, so that if the true mass in the above example was 3.20 g then these data are precise but inaccurate.

Quantitatively, a measure of precision would be a measure of spread of measured values. A measured mass of 2.7 g ± 0.1 g is less precise than 2.702 g ± 0.001 g. A quantitative treatment of precision is beyond the scope of the *VCE Chemistry Study Design*.

Replication of procedures: repeatability (reliability) and reproducibility

Experimental data and results must be more than one-off findings and should be repeatable and reproducible to draw reasonable conclusions. Repeatability refers to the closeness of agreement between independent results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus and/or same laboratory). Reproducibility refers to the closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus and/or different laboratories). The purposes of reproducing experiments include checking of claimed precision and uncovering of any systematic errors from one or other experiments/groups that may affect accuracy. Experiments that use subjective human judgment(s) or that involve small sample sizes or insufficient trials may also yield results that may not be repeatable and/or reproducible.

Validity

A measurement is 'valid' if it measures what it claims to be measuring. Both experimental design and the implementation should be considered when evaluating validity. Data are said to be valid if the measurements that have been made are affected by a single independent variable only. They are not valid if the investigation is flawed and control variables have been allowed to change or there is observer bias.

Experimental uncertainty and error

It is important not to confuse the terms 'error' and 'uncertainty', which are not synonyms. It is also important not to confuse 'error' with 'mistake' or 'personal error'. Error, from a scientific measurement perspective, is the difference between the measured value and the true value of what is being measured. Uncertainty is a quantification of the doubt associated with the measurement result. The *VCE Chemistry Study Design* requires only a qualitative treatment of uncertainty.

Experimental uncertainties are inherent in the measurement process and cannot be eliminated simply by repeating the experiment no matter how carefully it is done. There are two sources of experimental uncertainties: systematic effects and random effects. Experimental uncertainties are distinct from personal errors.

Personal errors

Personal errors include mistakes or miscalculations such as measuring a height when the depth should have been measured, or misreading the scale on a thermometer, or measuring the voltage across the wrong section of an electric circuit, or forgetting to divide the diameter by two before calculating the area of a circle using the formula $A = \pi r^2$. Personal errors can be eliminated by performing the experiment again correctly the next time, and do not form part of an analysis of uncertainties.

Systematic errors

Systematic errors are errors that affect the accuracy of a measurement. Systematic errors cause readings to differ from the true value by a consistent amount each time a measurement is made, so that all the readings are shifted in one direction from the true value. The accuracy of measurements subject to systematic errors cannot be improved by repeating those measurements.

Common sources of systematic errors include: faulty calibration of measuring instruments (and uncalibrated instruments) that consistently give the same inaccurate reading for the same value being measured, poorly maintained instruments (which may also have high random errors), or faulty reading of instruments by the user (for example, ‘parallax error’).

Random errors

Random errors affect the precision of a measurement and are always present in measurements (except for ‘counting’ measurements). These types of errors are unpredictable variations in the measurement process and result in a spread of readings.

Common sources of random errors are variations in estimating a quantity that lies between the graduations (lines) on a measuring instrument, the inability to read an instrument because the reading fluctuates during the measurement, and making a quick judgment of a transient event, for example, measuring the temperature at which a crystal first forms as a solution cools in order to construct a solubility curve.

The effect of random errors can be reduced by making more or repeated measurements and calculating a new mean and/or by refining the measurement method or technique.

Significant figures

Non-zero digits in data are always considered significant. Leading zeros are never significant whereas following zeros and zeros between non-zero digits are always significant. For example, 075.0210 contains six significant figures with the zero at the beginning not considered significant. A whole number may be a counting number or a measurement and determination of significant figures varies in the literature. For the purpose of the *VCE Chemistry Study Design*, whole numbers will have the same significant figures as number of digits, for example 400 has three significant figures while 400.0 has four.

Using a significant figures approach, one can infer the claimed accuracy of a value. For example, 400 is closer to 400 than 399 or 401. Similarly 0.0675 is closer to 0.0675 than 0.0674 or 0.0676.

Columns of data in tables should have the same number of decimal places, for example, measurements of lengths in centimetres or time intervals in seconds may yield the following data: 5.6, 9.2, 11.2 and 14.5. Significant figure rules should then be applied in subsequent data analysis.

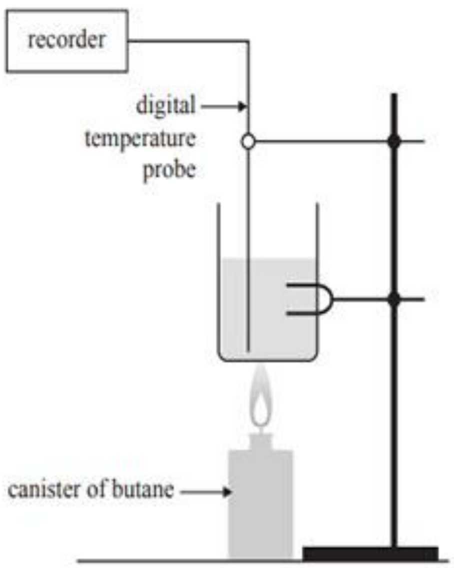
Calculations in chemistry often involve numbers having different numbers of significant figures. In mathematical operations involving:

- addition and subtraction, the student should retain as many digits to the right of the decimal as in the number with the fewest significant digits to the right of the decimal, for example: 386.38 + 793.354 - 0.000397 = 1179.73
- multiplication and division, the student should retain as many significant digits as in the number with the fewest significant digits, for example: 326.95 x 10.2 ÷ 20.322 = 164.

Intermediate results in calculations should retain at least one significant figure more than such analysis suggests until the final result is ascertained.

1) The molar heat of combustion of butane was calculated using the apparatus and the procedure shown below. Three trials were conducted and the results shown in the table. Label the following statements true or false. Explain why.

For this experiment, the students could maximise the:
A. precision by not using a digital thermometer ± 0.2 °C but rather using a manual mercury thermometer with a range of 20 -100°c.
B. validity by calculating the heat of combustion per mole.
C. accuracy by taking samples from three different sources.
D. uncertainty by having all students closely follow the same experimental procedure.



The heat content of butane

1. Measure the initial mass of a butane canister.
2. Measure the mass of a metal can, add 250 mL of water and re-weigh.
3. Set up the apparatus as in the diagram and measure the initial temperature of the water.
4. Burn the butane gas for five minutes.
5. Immediately measure the final temperature of the water.
6. Measure the final mass of the butane canister when cool.

Results table

	Trial 1	Trial 2	Trial 3
Quantity	**Measurement**	**Measurement**	**Measurement**
mass of empty can	52.14 g	52.45 g	52.35 g
mass of can + water before combustion	303.37 g	302.92 g	302.12 g
mass of butane canister before heating	260.15 g	268.25 g	251.89 g
mass of butane canister after heating	259.79 g	267.15 g	249.45 g
initial temperature of water	22.1 °C	22.5 °C	25.45°C
final temperature of water	32.7 °C	32.45 °C	32.5 °C

2) What is the difference between accuracy and validity? Explain using the experimental information given above.
3) For the results above.
   a) Calculate the heat of combustion of butane in kJ/g.
   a) Are the results accurate?
   b) Are the results repeatable or reliable?

c) What modifications can be made to get valid and Repeatable results.

4) What is the difference between random and systematic errors? How can each type be reduced?

5) An experiment is designed that properly addresses the aim, and care is taken with almost every factor. Repeating this experiment multiple times only leads to the same results. However, for some unknown reason, a systematic error has occurred that produces a result that is far from the result given in the literature. Another group in a different laboratory and time followed the same procedure and obtained the same value.
Explain why or why not the results are:
- valid
- accurate
- Repeatable/reliable
- reproducible

6) An experiment is designed that properly addresses the aim, while taking into account every variable. It is performed, once only, in a way that the results agree closely with the literature.
Explain why or why not the results:
- valid
- accurate
- repeatable/reliable
- reproducible

7) A student designs an experiment to measure the molar mass of CO_2 at SLC. To do this, the student designed a poor experiment where a number of variables were not controlled for. After several trials were conducted by the student and her peers a consistent value of around 44.0 g/mol was obtained for each trial.
Explain why or why not the results are:
- Valid
- Accurate
- Repeatable/reliable
- reproducible

8) The same student, as in 7) above, designs an experiment to measure the molar mass of CO_2 at SLC. To do this, the student designed a poor experiment where a number of variables were not controlled for. After several trials were conducted a consistent value of 66.0 g/mol was obtained for each trial.
Explain why or why not this experiment yields results that are:
- Valid
- Accurate
- Repeatable/reliable
- reproducible .

9) How is Repeatability and Validity related?

10) Joe designs an investigation, with many uncontrolled variables, to calculate the molar mass of propane. He conducts three trials and concludes the molar mass to be 44.0g/mol, 43.9g/mol and 44.1g/mol for each trial. Another group, following the exact method used by Joe, conducted the same experiment two days later and obtained the results below for the three trials.

40.0g/mol, 40.9g/mol and 39.1g/mol

Explain why or why not Joe's experiment yields results that are:

- valid
- accurate
- repeatable/reliable
- reproducible .

11) Joe designs a well structured investigation to calculate the molar mass of propane, with all variables, other than the dependent and independent variables, controlled. He conducts two sets of three trials and concludes the molar mass to be 44.0g/mol in the first set of trials and 45.0 g/mol in the second set. The results are shown below.

Set 1 43.9 g/mol, 44.1 g/mol, 44.0 g/mol

Set 2 45.2 g/mol, 45.1 g/mol, 45.0 g/mol

a) Explain why or why not the results are:
   - valid
   - accurate
   - repeatable/reliable
   - reproducible

b) What type of error occurred during the three trials in set 2? Explain and give an example.

12) Joe designs an investigation, with many uncontrolled variables, to calculate the molar mass of propane. He conducts three trials, averages the results and concludes the molar mass to be 44.0g/. The results are shown below.

41.9 g/mol, 47.1 g/mol, 43.0 g/mol

Indicate if the results are:

- valid
- accurate
- repeatable/reliable
- reproducible.

13) Consider the graph of reliability/repeatability versus accuracy shown below. Four different experiments were conducted. Which experiment/s:

a) minimised all errors and produced results that were very close to the true value?

b) was impacted by random errors?

c) was impacted by a systematic error?

d) "D" is both unreliable and yet accurate? Explain why

e) is impacted by errors that can be minimised through repetition ? Explain