Chemistry Units 3 & 4

Chapter 13 — Scientific investigations and data analysis

Aboriginal and Torres Strait Islander content has been excluded as accuracy cannot be guaranteed, and it is better to be respectful than cover the entire curriculum inaccurately.

Significant effort has been made to ensure this content is legal. Please send any areas of concern to admin@vwx.au with appropriate document/legal references so errors can be verified and changes be made.

Contents

Section	Page
13A Investigation design	2
13B Generating and evaluating evidence	12
13C Communicating chemistry	21

Chapter 13 Scientific investigations and data analysis — 13A Investigation design

Theory

Every chemical investigation begins with a plan. Before any data are collected, the investigator must decide *what* is being asked, *how* it will be tested, *which* quantities will be changed and measured, and *how* the quality and safety of the work will be assured. This subtopic sets out the design skills — aims and hypotheses, methodology and method, variables, data-generation techniques, measurement quality and safety — that underpin the scientific poster required by Unit 4 and that are examined in every end-of-year paper.

Aims, questions, hypotheses and predictions. An **aim** is a concise statement of the purpose of the investigation — what the experiment sets out to find. It usually names both the quantity being changed and the quantity being measured, for example "to investigate how the concentration of hydrochloric acid affects the rate of its reaction with marble chips". A **research question** poses the same idea as a question. A **hypothesis** is a testable, tentative prediction of the relationship between the two quantities, based on prior chemical knowledge; it is often written in an "if … then … because …" form and states a *direction* (increase or decrease). A **prediction** is the specific, observable outcome expected if the hypothesis is correct ("the powdered metal will react faster than the lump"). Designing a good investigation also requires identifying the **chemical concepts and key terms** it depends on — such as *collision theory*, *activation energy* or *standard electrode potential* — and defining them, so that the aim and hypothesis rest on sound chemistry.

Scientific methodology and method. The **methodology** is the overall *type* of scientific investigation chosen. In chemistry the usual choice is a **controlled experiment**, in which one variable is deliberately changed while all others are held constant; but the study design also recognises other methodologies — a *case study*, *classification and identification*, a *literature review*, *modelling*, *simulation*, *fieldwork*, and *product, process or system development*. The **method** is the specific, step-by-step procedure: the equipment, quantities, measurements and sequence of steps. A well-designed method is written in enough detail — specifying every quantity, instrument and condition — that the investigator can *repeat* it under the same conditions and another worker can *reproduce* it elsewhere and obtain comparable results; and it is **valid**, meaning it actually tests the stated aim.

Variables. In a controlled experiment there are three kinds of variable. The **independent variable (IV)** is the single quantity the investigator deliberately changes (for example, temperature, or concentration). The **dependent variable (DV)** is the quantity measured in response to that change (for example, the time for a precipitate to form, or a cell voltage). The **controlled variables** are all the other quantities that could affect the result and that are therefore kept constant from trial to trial. A sound design has exactly *one* IV and *one* DV; keeping every controlled variable constant is precisely what allows any change in the DV to be attributed to the IV, and is therefore the basis of a **valid** experiment.

Techniques of primary quantitative data generation. *Primary* data are generated first-hand by the investigator (as opposed to *secondary* data drawn from other sources). *Quantitative* data are numerical. The technique chosen must suit the quantity being measured, and the U3&4 course provides several, met in earlier chapters:

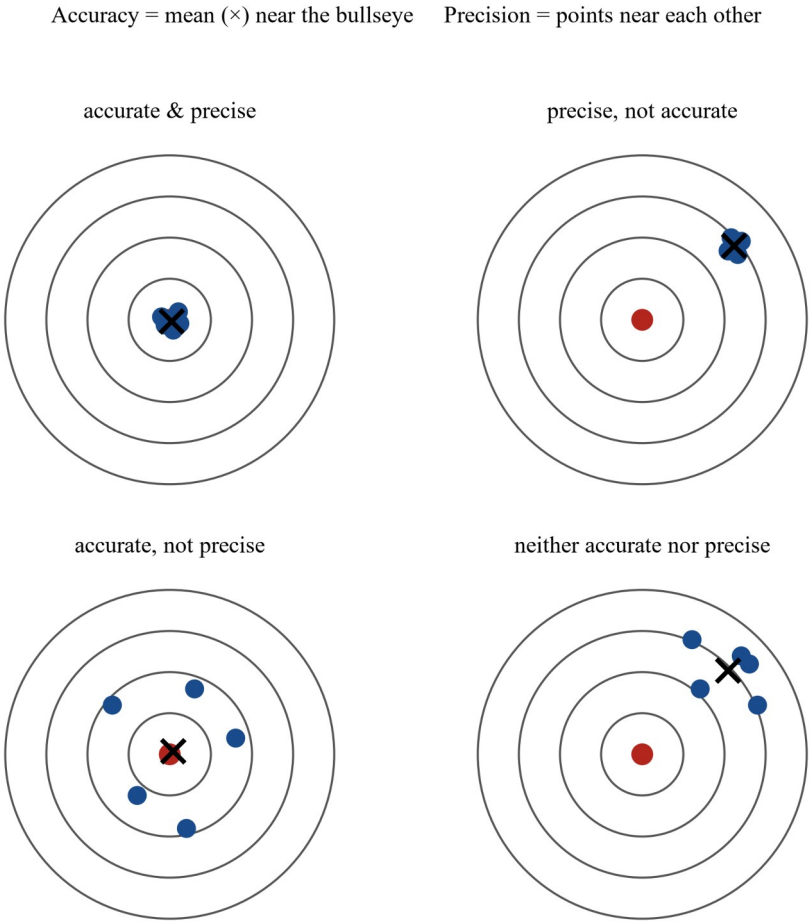
- **titration (volumetric analysis)** — to find the concentration or amount of a substance in solution (Chapter 10);
- **calorimetry** — to measure a heat or enthalpy change from a temperature change, $q = m\,c\,\Delta T$ (Chapter 2);
- **gas collection** — to follow the amount or rate of a reaction that produces a gas, by measuring gas volume against time (Chapter 5);
- **high-performance liquid chromatography (HPLC)** — to find the concentration of one component of a mixture from the area of its peak, read against a calibration curve (Chapter 11).

Measurement quality: accuracy, precision, repeatability, reproducibility and validity. These five terms are distinct and are frequently confused in examinations.

- **Accuracy** is how close a measurement (or, better, the *mean* of repeated measurements) is to the true or accepted value.
- **Precision** is how close repeated measurements are *to one another* — a small spread means high precision. Precision says nothing about accuracy: a measurement can be precise but wrong.

- **Repeatability** is the precision obtained when the *same* operator repeats the measurement with the *same* equipment and method, under the same conditions, over a short period.
- **Reproducibility** is the precision obtained when the measurement is repeated under *different* conditions — a different operator, different equipment, or a different laboratory.
- **Validity** describes whether the method actually measures what it is intended to measure — which requires an appropriate technique *and* proper control of variables. An experiment that lets a second variable change alongside the IV is not valid, however precise its readings.

The distinction between accuracy and precision is often shown on a target, where the bullseye is the true value, each dot is a measurement, and the cross marks the mean.



A tight cluster shows high precision; a mean (cross) near the bullseye shows high accuracy. A tight cluster *away* from the centre — precise but not accurate — is the signature of a **systematic error** (a consistent bias), whereas a wide scatter *around* the centre — accurate on average but not precise — is the signature of **random error**. (The full treatment of instrument resolution, of random and systematic error, and of uncertainty notation, belongs to subtopic 13B.)

Health, safety and ethical guidelines. Every practical design must be assessed for risk *before* it is carried out. A **risk assessment** identifies the hazards of each chemical and procedure — informed by the **safety data sheet (SDS)** for each substance — and states the precautions that reduce those risks: appropriate personal protective equipment (safety goggles, gloves, lab coat), working in a fume cupboard where toxic or volatile substances are handled, and safe techniques (for example, always adding concentrated acid to water, never the reverse). The design must also specify **safe disposal**: acidic or basic waste is neutralised before being flushed to the drain with plenty of water, while solutions of heavy-metal ions (such as copper(II) or lead(II)) are collected as chemical waste and never poured down the sink. **Ethical conduct** requires honest, complete recording of all results in a logbook — including anomalies — never altering or inventing data, and acknowledging all sources of information and assistance.

Worked examples

Worked Example 1 — A student investigates how the temperature of the reaction mixture affects the rate of reaction between sodium thiosulfate solution and dilute hydrochloric acid, $\mathrm{Na_2S_2O_3(aq) + 2\,HCl(aq) \rightarrow 2\,NaCl(aq) + S(s) + SO_2(g) + H_2O(l)}$. In each trial, 50.0 mL of $0.10\,\mathrm{mol\,L^{-1}}$ sodium thiosulfate and 5.0 mL of $2.0\,\mathrm{mol\,L^{-1}}$ hydrochloric acid are mixed in a conical flask standing on a printed cross, and the time for the sulfur precipitate to obscure the cross (viewed from above) is recorded, at 20, 30, 40 and 50 °C. State the aim, identify the independent, dependent and controlled variables, and write a hypothesis.

Working	Comment
Aim: to investigate how the temperature of the mixture affects the rate of reaction between sodium thiosulfate and hydrochloric acid.	The aim names the quantity changed (temperature) and the quantity measured (rate).
Independent variable: the temperature of the mixture (20 –50 °C).	The one quantity deliberately changed.
Dependent variable: the time for the cross to disappear.	Measured in response; a shorter time means a faster rate.
Controlled: volume and concentration of thiosulfate (50.0 mL, $0.10\,\mathrm{mol\,L^{-1}}$); volume and concentration of HCl (5.0 mL, $2.0\,\mathrm{mol\,L^{-1}}$); the same flask, cross, observer and viewing depth.	Held constant so the change in rate can be attributed to temperature alone — the basis of validity.
Hypothesis: if the temperature is increased, then the time for the cross to disappear will decrease (the rate will increase), because at higher temperature a greater proportion of collisions have energy exceeding the activation energy.	Testable and directional; links the IV to the DV with a chemical reason.

Worked Example 2 — A commercial vinegar is analysed for the concentration of ethanoic acid by titration; the accepted value is $0.100\,\mathrm{mol\,L^{-1}}$. Two students, working independently with different glassware, each perform three trials, obtaining the concentrations below. For each student, calculate the mean and comment on the precision and the accuracy; then state which measurement-quality term is tested by comparing the two students' results. Student P: 0.108, 0.109, 0.107 $\mathrm{mol\,L^{-1}}$. Student Q: 0.095, 0.104, 0.101 $\mathrm{mol\,L^{-1}}$.

Working	Comment
$\mathrm{Mean}_P = \dfrac{0.108 + 0.109 + 0.107}{3} = 0.108\,\mathrm{mol\,L^{-1}}$	Average of Student P's three trials.
$\mathrm{Range}_P = 0.109 - 0.107 = 0.002\,\mathrm{mol\,L^{-1}}$ (small) → **precise**	Small spread ⇒ high precision.
$\mathrm{Accuracy}_P$: 0.108 vs 0.100 is 8.0 % high → **not accurate**	A consistent bias — the mark of a systematic error.
$\mathrm{Mean}_Q = \dfrac{0.095 + 0.104 + 0.101}{3} = 0.100\,\mathrm{mol\,L^{-1}}$	Average of Student Q's three trials.
$\mathrm{Range}_Q = 0.104 - 0.095 = 0.009\,\mathrm{mol\,L^{-1}}$ (larger) → **less precise**	Wider spread ⇒ lower precision (random error).
$\mathrm{Accuracy}_Q$: mean = 0.100 = accepted → **accurate** (on average)	Close to the true value despite the scatter.
Comparing P's and Q's results tests **reproducibility**; each student's own three trials test **repeatability**.	Changed conditions (different operator, different glassware) ⇒ reproducibility.

Worked Example 3 — An investigation examines how the concentration of copper(II) ions affects the voltage of a galvanic cell. A $\mathrm{Cu^{2+}/Cu}$ half-cell is connected through a salt bridge to a standard $\mathrm{Zn^{2+}/Zn}$ half-cell, and the voltage is read on a digital voltmeter. The $[\mathrm{Cu^{2+}}]$ is set to 0.010, 0.10 and $1.0\,\mathrm{mol\,L^{-1}}$ in successive runs, while the

zinc half-cell is held at $1.0\,\text{mol L}^{-1}\,Zn^{2+}$ at 25 °C. Identify the variables, name the methodology, and explain how controlling the variables makes the investigation valid.

The **independent variable** is the concentration of copper(II) ions, $[Cu^{2+}]$ (0.010–$1.0\,\text{mol L}^{-1}$). The **dependent variable** is the measured cell voltage. The **controlled variables** are the zinc half-cell concentration ($1.0\,\text{mol L}^{-1}\,Zn^{2+}$), the temperature (25 °C), the identity and surface area of both electrodes, the salt-bridge electrolyte, and the volume of each solution. The **methodology** is a **controlled experiment**: one variable is deliberately changed while every other is held constant.

Because only $[Cu^{2+}]$ is changed while every other variable is held constant, any change in the measured voltage can be attributed to $[Cu^{2+}]$ alone; the experiment therefore measures what it is intended to measure. If, for instance, the temperature were allowed to drift between runs, a change in voltage could no longer be attributed to concentration, and the experiment would not be valid.

∴ IV = $[Cu^{2+}]$; DV = cell voltage; controlled = $[Zn^{2+}]$, temperature, electrodes, salt bridge and solution volumes; methodology = controlled experiment; and holding the controlled variables constant is exactly what makes the result a valid test of the effect of $[Cu^{2+}]$.

Worked Example 4 — Name the scientific methodology best suited to each of the following investigations, and give a reason. (a) Determining how the concentration of ethanoic acid affects the pH of the solution, with temperature and volume held constant. (b) Using a sequence of qualitative chemical tests to sort four unlabelled organic liquids into an alkanol, an alkanoic acid, an alkene and an ester. (c) Collating and analysing the results reported in five published journal articles on the efficiency of methanol fuel cells. (d) Designing, building and testing a prototype hand-warmer based on the crystallisation of sodium ethanoate. (e) For investigation (a), distinguish repeatability from reproducibility.

(a) A **controlled experiment**: one variable — the concentration of ethanoic acid — is deliberately changed while every other quantity that could affect the pH is held constant, so any change in pH can be attributed to concentration alone.

(b) **Classification and identification**: the investigation arranges substances into groups according to characteristic chemical behaviour and then names each one; no variable is manipulated.

(c) A **literature review**: the data are *secondary*, collated from the published work of others rather than generated first-hand, and the investigation consists of analysing and summarising them.

(d) **Product, process or system development**: an artefact is designed against an identified need, built, and then evaluated against performance criteria such as peak temperature, duration of warmth and reusability.

(e) **Repeatability** would be tested by the same student remeasuring the pH of the same solutions with the same pH meter and method, under the same conditions, over a short period; **reproducibility** would be tested by a different student, using a different pH meter or working in a different laboratory, obtaining the same relationship.

∴ (a) controlled experiment; (b) classification and identification; (c) literature review; (d) product, process or system development; and repeatability concerns the same operator, equipment and conditions, whereas reproducibility concerns changed conditions — a different operator, instrument or laboratory.

Practice questions

Question 1 — A student investigates how the concentration of hydrochloric acid affects the rate at which it reacts with excess marble chips (calcium carbonate). In each trial the same mass of marble chips is added to 50.0 mL of acid, and the volume of carbon dioxide collected in the first 30 seconds is measured. Concentrations of 0.50, 1.0, 1.5 and $2.0\,\text{mol L}^{-1}$ are used.

a. State the independent variable.
b. State the dependent variable.
c. State two variables that must be controlled.

Question 2 — A student plans to investigate how the surface area of zinc affects the rate at which it reacts with dilute sulfuric acid, comparing a single lump of zinc with the same mass of powdered zinc and measuring the volume of hydrogen gas produced per minute.

a. Write a suitable aim.
b. Write a hypothesis.
c. Write a prediction of what will be observed if the hypothesis is correct.

Question 3 — In standardising a solution, a student records three concordant titres: 24.30, 24.35 and 24.25 mL. The true titre for the standard is 24.30 mL.

a. Calculate the mean titre.
b. Using the spread of the titres, comment on the precision.
c. Comment on the accuracy of the student's titres.

Question 4 — A student investigates how the mass of sodium hydrogencarbonate added to a fixed volume of citric acid solution affects the volume of carbon dioxide produced in 60 s. For each trial they record the volume of gas collected and also write a description of how vigorously the mixture fizzes. Before starting, they read two published articles reporting this reaction.

a. Classify the gas volumes and the descriptions of the fizzing as quantitative or qualitative data.
b. Are the data in the two published articles primary or secondary data for this student? Explain.
c. Explain why the descriptive observations, although not numerical, should still be recorded in the logbook.

Question 5 — During a titration investigation the result is checked in two ways: (i) the same student repeats the titration three more times with the same burette and pipette; (ii) a second student in another class repeats the titration with different glassware and reports their result.

a. Which check tests the repeatability of the result?
b. Which check tests the reproducibility of the result?
c. Explain why obtaining a similar result in both checks strengthens the conclusion.

Question 6 — A student designs a controlled experiment to measure how the initial temperature of hydrogen peroxide affects the rate of its catalysed decomposition, $2\,H_2O_2(aq) \rightarrow 2\,H_2O(l) + O_2(g)$, by collecting the oxygen produced. In their first attempt they use a different mass of manganese(IV) oxide catalyst in each trial.

a. Name the scientific methodology being used.
b. Identify the flaw that makes the first attempt invalid, and explain how to correct it.
c. State one safety precaution and one appropriate way to dispose of the reaction mixture.

Question 7 — In a calorimetry investigation, a student measures how the concentration of hydrochloric acid affects the temperature rise when it is neutralised by an excess of sodium hydroxide. Equal volumes (50.0 mL) of acid at 0.50, 1.0 and $2.0\,\text{mol L}^{-1}$ are each added to 50.0 mL of $2.5\,\text{mol L}^{-1}$ sodium hydroxide in an insulated cup, and the maximum temperature rise is recorded. Identify the independent, dependent and controlled variables, and write a balanced equation, with states, for the neutralisation reaction.

Question 8 — A student will build galvanic cells pairing a standard zinc half-cell with the half-cell of a second metal (magnesium, iron, copper or silver) and measure the cell voltage, to investigate how the second metal affects the voltage. Write an aim and a hypothesis, and state the independent and dependent variables.

Question 9 — Two students each determine the mass of copper deposited in an electrolysis experiment three times; the true mass is 5.00 g. Student A records 5.55, 5.56 and 5.54 g; Student B records 4.90, 5.05 and 5.05 g. For each student, calculate the mean and the range, describe the results in terms of accuracy and precision, and suggest the type of error affecting Student A.

Question 10 — For an investigation into how temperature affects the solubility of potassium nitrate, classify each of the following statements as an aim, a hypothesis or a prediction: (a) "At 60 °C, more than 100 g of potassium nitrate will dissolve in 100 g of water."; (b) "To investigate how the temperature of the water affects the mass of potassium

nitrate that dissolves in 100 g of water."; (c) "If the temperature of the water is increased, then the mass of potassium nitrate that dissolves will increase, because dissolving this salt is endothermic and so is favoured at higher temperature."

Question 11 — A research team publishes a new method for measuring the vitamin C content of juice by redox titration. Explain, using the terms *repeatability* and *reproducibility*, what two kinds of check would show that the method is reliable.

Question 12 — A student investigates how the concentration of hydrochloric acid affects the rate of its reaction with magnesium ribbon, but carries out the higher-concentration trials in the afternoon, when the laboratory is several degrees warmer than in the morning. Explain why this threatens the validity of the investigation, and state how the method should be changed.

Question 13 — A titration uses concentrated sulfuric acid to acidify the analyte. Referring to a risk assessment, state one hazard of concentrated sulfuric acid, one precaution informed by its safety data sheet (SDS), and how the neutralised waste should be disposed of.

Question 14 — A student measures the enthalpy of combustion of ethanol by burning it in a spirit burner under an open metal can of water, records the temperature rise once, and obtains a value much smaller in magnitude than the accepted value. Evaluate the method and suggest three improvements that would give a result closer to the accepted value.

Question 15 — For each investigation, name the most appropriate technique of primary quantitative data generation: (a) finding the concentration of a sodium hydroxide solution; (b) measuring the molar enthalpy change of a neutralisation reaction; (c) following the rate of a reaction that produces a gas; (d) measuring the concentration of caffeine in a sample of energy drink.

Question 16 — A student records five titres from a burette graduated at every 0.1 mL: 25.10, 25.15, 25.10, 25.90 and 25.15 mL.

a. Identify the outlier.
b. Calculate the mean of the concordant titres (excluding the outlier).
c. State the range of the concordant titres and comment on the precision.

Question 17 — Explain why, in an investigation of how temperature affects the rate of a reaction, it is important to keep the concentration and volume of the reactants the same in every trial. Use the term *valid* in your answer.

Question 18 — A student is given the research question: "How does the concentration of a copper(II) sulfate solution affect the mass of copper deposited on the cathode during electrolysis in a fixed time?" Design the investigation by stating (a) an aim, (b) a hypothesis, (c) the independent variable, (d) the dependent variable, (e) three controlled variables, (f) how the primary quantitative data will be generated, and (g) one health-and-safety consideration.

Answers

1 a the concentration of the hydrochloric acid; b the volume of CO_2 collected in 30 s (a measure of rate); c any two of: mass/surface area of marble chips, volume of acid (50.0 mL), temperature, the fixed 30 s interval. **2** a to investigate how the surface area of zinc affects the rate of its reaction with dilute sulfuric acid; b if the surface area is increased, then the rate of hydrogen production increases, because more zinc atoms are exposed so collisions are more frequent; c the powdered zinc produces hydrogen faster (a greater volume per minute) than the lump of equal mass. **3** a $24.30\,\text{mL}$; b range $= 0.10\,\text{mL}$, a small spread ⇒ precise (concordant); c the mean equals the true value ⇒ accurate. **4** a the gas volumes are quantitative (numerical); the descriptions of the fizzing are qualitative (non-numerical); b secondary — they were generated by other investigators, not first-hand by this student (only the student's own gas volumes are primary); c a logbook must be a complete, honest record of everything observed, including anomalies; a qualitative note often explains a volume that does not fit the trend and lets another chemist scrutinise or reproduce the work. **5** a check (i) — same student, same equipment; b check (ii) — different student, different equipment; c agreement under both the same and different conditions shows the result is consistent and not an artefact of one operator or apparatus, so the evidence is more reliable. **6** a a controlled experiment; b the catalyst mass is not controlled — it changes alongside temperature, so any change in rate cannot be attributed to temperature alone (not valid); use the same mass of catalyst in every trial; c e.g. wear safety goggles (keep the O_2-releasing mixture away from flames); filter off the manganese(IV) oxide as solid waste and rinse the dilute solution down the sink with water. **7** IV = concentration of HCl; DV = maximum temperature rise; controlled = volume of acid (50.0 mL), volume and concentration of NaOH (50.0 mL, $2.5\,\text{mol L}^{-1}$), the insulated calorimeter, initial temperature; $\text{HCl(aq)} + \text{NaOH(aq)} \rightarrow \text{NaCl(aq)} + \text{H}_2\text{O(l)}$. **8** aim: to investigate how the identity of the second metal (paired with a standard zinc half-cell) affects the cell voltage; hypothesis: if the second metal is a weaker reductant (higher standard reduction potential) than zinc, then the cell voltage is larger, because $E^\circ_{\text{cell}} = E^\circ(\text{cathode}) - E^\circ(\text{anode})$ is greater; IV = identity of the second metal, DV = cell voltage. **9** A: mean $= 5.55\,\text{g}$, range $= 0.02\,\text{g}$ ⇒ precise but not accurate (systematic error); B: mean $= 5.00\,\text{g}$, range $= 0.15\,\text{g}$ ⇒ accurate but not precise (random error). **10** a prediction; b aim; c hypothesis. **11** repeatability — the same analyst repeats the titration with the same equipment and obtains concordant results; reproducibility — a different analyst or laboratory, using different equipment, obtains a similar result. A reliable method shows both. **12** temperature also affects rate, and it is not controlled (it changes with concentration), so a change in rate could be due to the warmer afternoon rather than the concentration — the experiment is not a valid test of concentration; carry out all trials at the same temperature. **13** hazard: it is highly corrosive, causing severe burns to skin and eyes; precaution (per SDS): wear goggles, gloves and a lab coat, work in a fume cupboard, and add acid to water; disposal: neutralise the waste (e.g. with sodium hydrogencarbonate), then dilute well and rinse down the sink. **14** the open apparatus loses heat to the surroundings, the can and the hot gases, so the measured ΔT (and $|\Delta H|$) is too small; improvements: insulate/shield the apparatus and reduce the flame-to-can distance; repeat and average; use an electrically calibrated calorimeter. **15** a titration (volumetric analysis); b calorimetry; c gas collection (gas volume vs time); d HPLC, read against a calibration curve. **16** a $25.90\,\text{mL}$; b mean $= 25.125 \approx 25.13\,\text{mL}$; c range $= 0.05\,\text{mL}$, a very small spread ⇒ precise. **17** concentration and volume also affect the rate; if they changed between trials, a change in rate could not be attributed to temperature; keeping them constant makes temperature the only variable that changes, so the experiment is a valid test of temperature's effect. **18** a to investigate how $[\text{CuSO}_4]$ affects the mass of copper deposited on the cathode in a fixed time; b if $[\text{CuSO}_4]$ increases, then the mass deposited increases, because a more concentrated solution conducts better and supplies Cu^{2+} to the cathode faster, so at the fixed applied voltage a larger current flows; c IV = $[\text{CuSO}_4]$; d DV = mass of copper deposited; e any three of: applied voltage, time, electrode area/material, electrode separation, solution volume, temperature; f weigh the dried cathode on an electronic balance before and after; the increase in mass is the mass of copper deposited; g e.g. copper(II) sulfate is harmful and toxic to aquatic life — wear goggles and collect it as heavy-metal waste, not down the drain.

Fully-worked solutions

Question 1 a. The **independent variable** is the concentration of the hydrochloric acid — the quantity deliberately changed (0.50–$2.0\,\text{mol L}^{-1}$). b. The **dependent variable** is the volume of carbon dioxide collected in the first 30 seconds, which measures how fast the reaction proceeds. c. Any two of: the mass (and surface area) of the marble chips, the volume of acid (50.0 mL), the temperature, and the fixed 30 s collection time. Keeping these constant

ensures that only the acid concentration varies. (The reaction is $\mathrm{CaCO_3(s) + 2\,HCl(aq) \rightarrow CaCl_2(aq) + H_2O(l) + CO_2(g)}$.)

Question 2 a. Aim: to investigate how the surface area of zinc affects the rate of its reaction with dilute sulfuric acid. b. Hypothesis: if the surface area of the zinc is increased (by powdering it), then the rate of hydrogen production will increase, because a greater surface area exposes more zinc atoms to the acid particles, making collisions more frequent. c. Prediction: the powdered zinc will react faster, producing a greater volume of hydrogen gas per minute than the single lump of the same mass. (The reaction is $\mathrm{Zn(s) + H_2SO_4(aq) \rightarrow ZnSO_4(aq) + H_2(g)}$.)

Question 3 a. Mean $= \dfrac{24.30 + 24.35 + 24.25}{3} = \dfrac{72.90}{3} = 24.30\,\mathrm{mL}$. b. Range $= 24.35 - 24.25 = 0.10\,\mathrm{mL}$. This is a small spread, so the titres are close to one another — they are **precise** (concordant). c. The mean, $24.30\,\mathrm{mL}$, is equal to the true value of $24.30\,\mathrm{mL}$, so the result is **accurate**. (A set can be both accurate and precise, as here.)

Question 4 a. The volumes of carbon dioxide are **quantitative** data — numerical measurements on a scale. The written descriptions of how vigorously the mixture fizzes are **qualitative** data — they describe the behaviour in words rather than as a number. b. **Secondary** data: they were generated by other investigators and obtained by this student from a published source. Only the measurements the student makes first-hand — the gas volumes from their own trials — are **primary** data. c. Ethical conduct requires an honest and complete record of everything observed, including anomalies. A qualitative note ("the bung was not properly seated in trial 3", "this mixture fizzed over the top of the flask") often explains a gas volume that does not fit the trend, and lets a reader — or a second investigator repeating the work — judge whether that result should be retained. Recording only the convenient numbers would misrepresent the investigation.

Question 5 a. Check (i) tests **repeatability**: the same student repeats the measurement with the same equipment and method. b. Check (ii) tests **reproducibility**: a different student repeats the measurement with different glassware. c. If the result is similar when the same person repeats it (repeatable) *and* when a different person with different equipment repeats it (reproducible), then the result is consistent and does not depend on one operator or one set of apparatus. This makes the evidence more reliable and the conclusion more trustworthy.

Question 6 a. A **controlled experiment** — one variable (the initial temperature) is deliberately changed while the others are held constant. b. The **mass of catalyst is not controlled**: it is changed from trial to trial alongside the temperature. A catalyst affects the rate, so any change in the rate could be caused by the differing catalyst mass rather than the temperature; the experiment therefore does not validly test the effect of temperature. To correct it, use the **same mass** of manganese(IV) oxide in every trial, changing only the temperature. c. Safety precaution: wear safety goggles (hydrogen peroxide is an irritant/corrosive and the oxygen produced supports combustion, so keep the apparatus away from naked flames). Disposal: filter off the manganese(IV) oxide as solid waste; the remaining dilute mixture is essentially water and may be rinsed down the sink with plenty of water.

Question 7 Independent variable: the concentration of hydrochloric acid (0.50–$2.0\,\mathrm{mol\,L^{-1}}$). Dependent variable: the maximum temperature rise of the mixture. Controlled variables: the volume of acid (50.0 mL), the volume and concentration of sodium hydroxide (50.0 mL, $2.5\,\mathrm{mol\,L^{-1}}$ — an excess, so that every trial's acid is fully neutralised), the same insulated calorimeter, and the initial temperature. Balanced equation: $\mathrm{HCl(aq) + NaOH(aq) \rightarrow NaCl(aq) + H_2O(l)}$.

Question 8 Aim: to investigate how the identity of the second metal, paired with a standard zinc half-cell, affects the voltage of the galvanic cell. Hypothesis: if the second metal is a weaker reductant than zinc (a higher standard reduction potential), then the cell voltage will be larger, because the cell voltage is the difference between the two electrode potentials, $E^\circ_{\text{cell}} = E^\circ(\text{cathode}) - E^\circ(\text{anode})$, which is greater when the second electrode sits further from zinc on the electrochemical series. Independent variable: the identity of the second metal (Mg, Fe, Cu or Ag). Dependent variable: the measured cell voltage.

Question 9 Student A: mean $= \dfrac{5.55 + 5.56 + 5.54}{3} = \dfrac{16.65}{3} = 5.55\,\mathrm{g}$; range $= 5.56 - 5.54 = 0.02\,\mathrm{g}$. The three values are close to one another (**precise**) but their mean, $5.55\,\mathrm{g}$, is well above the true value of $5.00\,\mathrm{g}$ (**not accurate**). A consistent bias of this kind indicates a **systematic error**. Student B: mean $= \dfrac{4.90 + 5.05 + 5.05}{3} = \dfrac{15.00}{3} = 5.00\,\mathrm{g}$;

range $= 5.05 - 4.90 = 0.15\,\text{g}$. The mean equals the true value (**accurate**) but the values are widely spread (**not precise**), which indicates **random error**.

Question 10 a. A **prediction** — a specific, observable outcome expected at one stated value of the independent variable (60 °C), against which the investigation can be checked. b. An **aim** — a concise statement of the purpose of the investigation, naming both the quantity deliberately changed (the temperature of the water) and the quantity measured (the mass of potassium nitrate that dissolves). c. A **hypothesis** — a testable, directional statement of the expected relationship between the IV and the DV, in "if … then … because …" form and supported by a chemical reason.

Question 11 Repeatability: the same analyst repeats the redox titration several times with the same equipment and method and obtains closely-agreeing (concordant) results, showing that the method is internally consistent. **Reproducibility**: a different analyst, or a different laboratory using different equipment, follows the same method and obtains a similar result, showing that the finding does not depend on one operator or one set of apparatus. A reliable method should demonstrate both.

Question 12 Temperature also affects the rate of a reaction. In this design the temperature is not controlled — it rises through the day and so changes together with the acid concentration (the higher-concentration trials are also the warmer ones). As a result, any increase in rate could be caused by the warmer afternoon rather than by the higher concentration; the two possible causes are confounded, so the experiment does not validly measure the effect of concentration alone. The method should be changed so that all trials are carried out at the **same, controlled temperature** — for example by thermostatting the reactants in a water bath, or by performing all trials together under identical conditions — leaving concentration as the only variable that changes.

Question 13 Hazard: concentrated sulfuric acid is highly **corrosive**, causing severe burns to skin and eyes, and it reacts strongly (exothermically) with water. Precaution (informed by the SDS): wear safety goggles, gloves and a lab coat; handle it in a fume cupboard; and when diluting, add the acid slowly to water (never water to acid) to control the heat released. Disposal: neutralise the acidic waste — for example with sodium hydrogencarbonate — until it is close to neutral, then dilute it with plenty of water and rinse it down the sink (or place it in the designated inorganic-waste container), as directed by the SDS.

Question 14 Evaluation: in the open "flame under a can" apparatus, much of the heat released by the burning ethanol is lost to the surrounding air, absorbed by the metal can, or carried away by the rising hot combustion gases, rather than passing into the water. The measured temperature rise is therefore smaller than it should be, so the calculated heat and the enthalpy magnitude come out too low. Taking only a single reading also gives no check on precision, and the combustion in an open flame may be incomplete. Three improvements: (1) reduce heat loss by insulating and shielding the apparatus (a draught shield and a lid) and by reducing the distance between the flame and the can; (2) repeat the measurement several times and average the results, to improve precision and reveal anomalies; (3) use an insulated calorimeter that has been **electrically calibrated**, so the heat capacity of the whole apparatus is accounted for. (Extrapolating a temperature–time graph to correct for heat loss would further improve the result.)

Question 15 a. **Titration (volumetric analysis)** — the concentration of a solution is found by reacting it with a standard solution to an endpoint. b. **Calorimetry** — a heat change is found from a measured temperature change, using $q = m\,c\,\Delta T$. c. **Gas collection** — the volume of gas produced is measured against time to follow the rate. d. **High-performance liquid chromatography (HPLC)** — the drink is separated on the column and the concentration of caffeine is read from the area of its peak against a calibration curve prepared from standards of known concentration.

Question 16 a. The **outlier** is $25.90\,\text{mL}$: it lies far from the other four readings, which are grouped near 25.1 mL. b. Excluding the outlier, mean $= \dfrac{25.10 + 25.15 + 25.10 + 25.15}{4} = \dfrac{100.50}{4} = 25.125 \approx 25.13\,\text{mL}$. (Each individual titre must sit on a graduation or halfway between two, i.e. on a multiple of $0.05\,\text{mL}$; a *mean*, being calculated rather than read, may carry a further digit, but is conventionally reported to the same two decimal places as the titres.) c. Range of the concordant titres $= 25.15 - 25.10 = 0.05\,\text{mL}$. This is a very small spread, so the concordant titres are **precise**.

Question 17 The concentration and the volume of the reactants both affect the rate of a reaction. If either of them were allowed to change between trials, then a change in the measured rate could be caused by that change rather than by the temperature under investigation. By keeping the concentration and volume constant (as **controlled variables**),

the temperature becomes the only quantity that varies, so any change in rate can be attributed to temperature alone. This is what makes the experiment **valid** — it ensures the investigation actually measures the effect of temperature, which is what it is intended to measure.

Question 18 a. Aim: to investigate how the concentration of a copper(II) sulfate solution affects the mass of copper deposited on the cathode during electrolysis in a fixed time. b. Hypothesis: if the concentration of copper(II) sulfate is increased, then the mass of copper deposited in the fixed time will increase, because a more concentrated solution conducts electricity better and delivers Cu^{2+} ions to the cathode faster, so at the fixed applied voltage a larger current flows and — by Faraday's laws — a greater mass of copper is reduced onto the cathode in that time. c. Independent variable: the concentration of the copper(II) sulfate solution. d. Dependent variable: the mass of copper deposited on the cathode. e. Controlled variables (any three): the applied voltage, the electrolysis time, the surface area and material of the electrodes, the distance between the electrodes, the volume of solution, and the temperature. (The current is deliberately *not* fixed here: it is the quantity through which the concentration acts, so fixing it as well would remove the effect being investigated.) f. The primary quantitative data are generated by **mass measurement**: the cathode is dried and weighed on an electronic balance before and after electrolysis, and the increase in mass is the mass of copper deposited. g. Health-and-safety consideration: copper(II) sulfate solution is harmful (an irritant) and toxic to aquatic life, so safety goggles should be worn and the solution collected as heavy-metal chemical waste rather than poured down the drain.

Chapter 13 Scientific investigations and data analysis — 13B Generating and evaluating evidence

Theory

An investigation is only as good as the **evidence** it generates and the way that evidence is analysed. This subtopic is about turning raw measurements into a defensible conclusion: organising and processing primary data, judging its quality, accounting for error and uncertainty, and stating honestly how far the evidence can be trusted. The skills apply to every quantitative context in Units 3 & 4 — titration, calorimetry, rates, electrochemistry and instrumental analysis.

The nature of scientific evidence. Evidence is data, generated by a valid method, that bears on a hypothesis, model or theory. It is not the same as **opinion** or **anecdote**: a single unrepeated observation, or a claim made without data, carries little evidential weight. Evidence can **support** a hypothesis (the results are consistent with it) or **refute** it (the results contradict it). Crucially, evidence never **proves** a hypothesis: a conclusion in science is always **provisional** and is **bounded by the data** that produced it. Concordant results make a hypothesis more strongly supported, but the true value can only ever be estimated within a range, never confirmed with certainty.

Organising and analysing primary data. Primary data is data you generate first-hand in your own investigation (as distinct from secondary data collected from other sources). Raw data is first organised into a **table** with clearly headed columns and units, then **processed** to reveal what it means. Common processing steps are calculating a **mean** of repeated measurements, a **ratio**, a **percentage** or **percentage change**, and identifying an overall **trend**. Presenting the data as a **graph** then exposes the **pattern or relationship** between the variables far more clearly than a table of numbers.

Graphs, lines of best fit and the meaning of a gradient. When two variables are plotted, a smooth **line (or curve) of best fit** is drawn so that the points are balanced evenly on either side — it is *not* a dot-to-dot join, because joining each point would treat the random scatter of individual points as if it were real. The line of best fit **averages out** that scatter. If the plotted points lie on a straight line, the two variables are **linearly related**; if that line passes through the origin, they are **directly proportional**. The **gradient** of the line has a physical meaning set by the axes: for a plot of gas volume against mass of reactant the gradient is "volume of gas produced per gram of reactant"; for a plot of volume of gas against time it is the **rate** of the reaction. Always quote a gradient **with its units** and interpret what it represents.

Repeats, concordant results and outliers. Repeating a measurement lets you judge — and improve — the quality of the data. Repeats that agree closely are **concordant** (for titres, agreeing to within about $0.10\,\text{mL}$). An **outlier** is a value that lies well outside the concordant group; it usually arises from a one-off mistake (an overshot end point, a gas leak, a misread scale). An outlier should be **identified and excluded, with a justification**, before the concordant results are averaged — otherwise it drags the mean away from the true value. A mean of the concordant results is therefore more **accurate** — closer to the true value — than a mean that includes an anomalous result.

Uncertainty and resolution. Every measurement carries an **uncertainty**, written with a ± symbol. The **resolution** of an instrument is the smallest interval it can distinguish. For an **analogue** graduated scale (a burette, a thermometer), a single reading is usually estimated to ± **half the smallest division**: a burette graduated every $0.1\,\text{mL}$ gives a reading uncertainty of $\pm 0.05\,\text{mL}$. For a **digital** instrument the resolution is the **last displayed digit**, and a reading is quoted to ± that last digit: a balance showing $2.47\,\text{g}$ has a resolution of $0.01\,\text{g}$ and its reading is $(2.47 \pm 0.01)\,\text{g}$. When a result is a **difference** of two readings (a titre = final − initial), each reading carries this uncertainty, so the titre is quoted to the same number of decimal places as the readings — $21.50\,\text{mL}$, never 21.5 and never 21.53.

VCE Chemistry requires only a **qualitative** treatment of uncertainty: you are expected to state the resolution of an instrument and the ± uncertainty of a single reading, but *not* to combine uncertainties arithmetically or to attach a calculated ± figure to a mean. Where a quantity has been measured several times, it is the **spread** of the repeats that is discussed: closely grouped repeats indicate a **precise** measurement result, a widely spread set an imprecise one.

Random and systematic error; accuracy and precision. Errors come in two kinds.

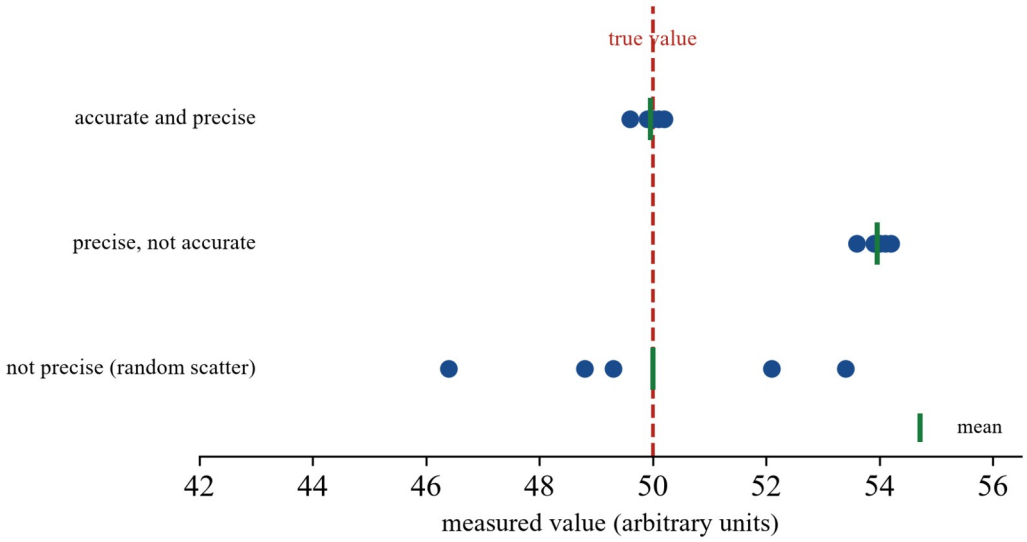
- A **random error** scatters readings unpredictably above and below the true value (estimating the last figure of a scale between two graduations, judging the exact shade of an indicator's end-point colour, timing variation

from trial to trial, small fluctuations in laboratory conditions). It reduces the **precision** of the data, and its effect is **reduced by taking repeats and averaging**.

- A **systematic error** shifts every reading in the **same direction** by a similar amount. It arises from a **limitation of the instrument** or its **calibration** (an un-zeroed balance, a mis-calibrated thermometer) or from an **inappropriate method** (a wrongly made-up standard, or **parallax** — reading a scale from an angle instead of at eye level). It reduces the **accuracy** of the data and is **not** reduced by repeating — averaging simply reproduces the same bias. A systematic error is found by **calibration** or by comparison with an accepted value.

Watch parallax. It is tempting to call parallax a random error, because different people read a scale differently. In VCE Chemistry it is classified as a **systematic** error: it is an inappropriate method, and it shifts a given observer's readings consistently in one direction. Reading every scale at eye level removes it; repeating and averaging do not.

These map onto two separate ideas: **accuracy** is how close a result is to the **true (accepted) value**; **precision** is how close the repeated results are to **each other**. Data can be precise but not accurate (tightly grouped but off-target — a systematic error), or accurate on average but imprecise (widely scattered about the true value — random error).



Authentication through a logbook. Primary data is **authenticated** by a **logbook**: a dated, contemporaneous record of the raw readings, observations and any changes made to the method, kept as the investigation proceeds and signed as the student's own work. The logbook is the evidence that the data is genuinely what was generated — not altered or invented afterwards — and it lets another chemist trace, check or reproduce the results.

Assumptions, limitations and bounding the conclusion. Every method rests on **assumptions** (for example, that a titration's indicator colour change coincides with the equivalence point, or that no heat is lost from a calorimeter) and has **limitations** in how the data is generated and analysed. A conclusion is only valid **within the conditions and range actually tested**: a relationship measured between 20 °C and 50 °C cannot be assumed to hold at 80 °C, and reading a value off the far end of a graph (extrapolation) assumes the pattern continues. Stating these assumptions and limitations honestly — and recognising that the conclusion is bounded by the data — is a required part of evaluating evidence.

Worked examples

Worked Example 1 — In a titration, a student records five titres, in mL: 25.60, 24.95, 25.05, 25.00, 25.00. Determine the mean titre that should be used, justifying any value you exclude.

Working	Comment
Titres/mL: 25.60, 24.95, 25.05, 25.00, 25.00	List the raw data; the first is a rough (trial) titre.
25.60 lies about 0.6 mL above the others	25.60 is an **outlier** — it is not concordant with the cluster 24.95–25.05 mL.
Concordant titres: 24.95, 25.05, 25.00, 25.00	Concordant titres agree to within about 0.10 mL.

Working	Comment
$\text{mean} = \dfrac{24.95 + 25.05 + 25.00 + 25.00}{4} = \dfrac{100.00}{4} = 25.00$	Average **only** the concordant titres.
(all five would give $125.60/5 = 25.12\ \text{mL}$)	Including the outlier biases the mean upward by $0.12\ \text{mL}$.

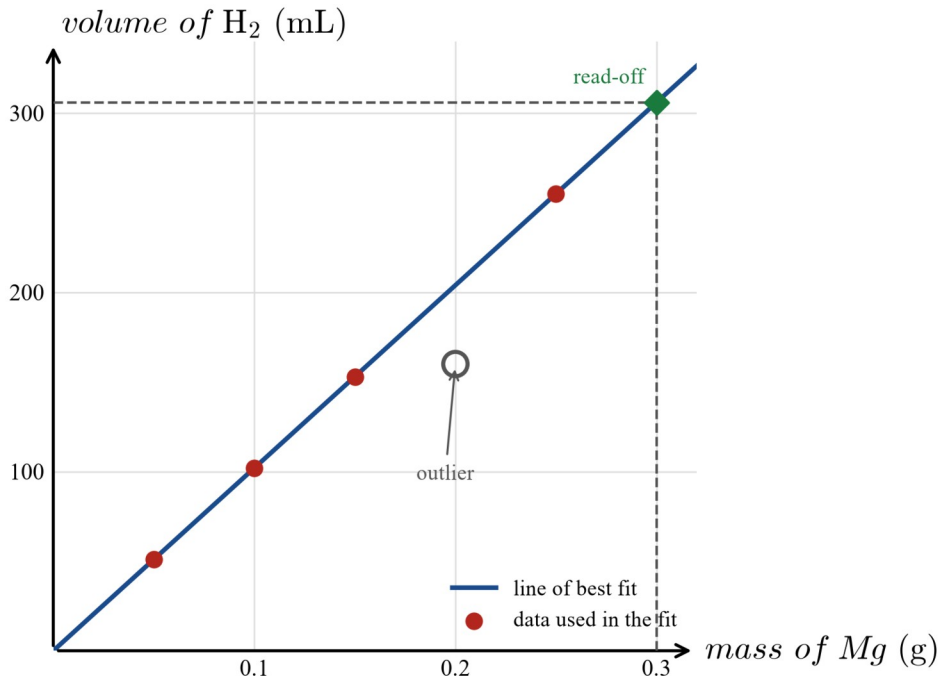
Worked Example 2 — A burette graduated in $0.1\ \text{mL}$ divisions gives an initial reading of $1.05\ \text{mL}$ and a final reading of $22.55\ \text{mL}$. (a) State the resolution of the burette, the uncertainty in each reading, and the titre delivered. (b) During the titration the student judged the exact shade of the indicator's end-point colour slightly differently from one trial to the next; separately, the pipette used was found to deliver slightly less than its nominal volume in every trial. Classify each error and state its effect.

Working	Comment
$\text{resolution} = 0.1\ \text{mL}$	The smallest interval the burette scale can distinguish.
$(1.05 \pm 0.05)\ \text{mL}$ and $(22.55 \pm 0.05)\ \text{mL}$	Analogue scale: a single reading is estimated to ± half the smallest division.
$\text{titre} = 22.55 - 1.05 = 21.50\ \text{mL}$	The titre is a **difference** of two readings, so it is quoted to the same two decimal places. Uncertainty is treated **qualitatively** — do not add the two ± figures together.
varying judgement of the end-point shade → **random** error	The judgement falls sometimes early, sometimes late; this lowers **precision** and is reduced by averaging repeats.
pipette under-delivering each trial → **systematic** error	The same bias, same direction, every time; this lowers **accuracy** and is **not** reduced by repeats.

Worked Example 3 — A student reacts different masses of magnesium ribbon with **excess** hydrochloric acid and measures the volume of hydrogen gas collected at SLC:

mass of Mg/g	0.05	0.10	0.15	0.20	0.25
$V(\text{H}_2)/\text{mL}$	51	102	153	160	255

Plot the data, draw a line of best fit, and use it to (a) find the gradient and state its meaning, and (b) predict the gas volume for $0.30\ \text{g}$ of magnesium.



The point $(0.20, 160)$ lies well below the trend of the other four points — it is an **outlier** (hydrogen most likely leaked past the bung), so it is excluded before fitting. No gas forms when no magnesium is added, so the line of best fit is drawn **through the origin**. Its gradient is

$$\text{gradient} = \frac{\Delta V}{\Delta m} = \frac{255 - 0}{0.25 - 0} = 1020\ \text{mL g}^{-1}$$

The gradient is the **volume of hydrogen produced per gram of magnesium**: each gram of magnesium yields about $1020\ \text{mL}$ (i.e. $1.02\ \text{L}$) of hydrogen at SLC. This agrees with the stoichiometry of $\text{Mg(s)} + 2\,\text{HCl(aq)} \rightarrow \text{MgCl}_2\text{(aq)} + \text{H}_2\text{(g)}$, in which $n(\text{H}_2) = n(\text{Mg})$: for $1.00\ \text{g}$ of magnesium, $n = \frac{1.00}{24.3} = 0.0412\ \text{mol}$, so $V = 0.0412 \times 24.8 = 1.02\ \text{L}$. Reading off the line at $m = 0.30\ \text{g}$ gives $V = 1020 \times 0.30 = 306\ \text{mL}$.

$\therefore$ the gradient is $1020\ \text{mL g}^{-1}$ and the predicted volume is $306\ \text{mL}$ — but this reading lies just beyond the measured range ($0.30 > 0.25\ \text{g}$), so it **assumes** the acid is still in excess and the proportional relationship continues to hold.

Worked Example 4 — A student's hypothesis is that a hydrochloric acid solution has a concentration of $0.100\ \text{mol L}^{-1}$. Three independent titrations give 0.098, 0.101 and $0.102\ \text{mol L}^{-1}$. (a) Determine the measurement result and comment on the precision of the three titrations. (b) Evaluate whether the evidence supports the hypothesis. (c) State one assumption of the analysis and explain how a logbook authenticates the data.

$$\text{measurement result} = \text{mean} = \frac{0.098 + 0.101 + 0.102}{3} = \frac{0.301}{3} = 0.100\ \text{mol L}^{-1}\ (3\,\text{sf})$$

The three values are spread over only 0.098–$0.102\ \text{mol L}^{-1}$, a spread of $0.004\ \text{mol L}^{-1}$ about the mean. They are therefore closely grouped and **concordant**, so the titrations are **precise**.

The hypothesised value $0.100\ \text{mol L}^{-1}$ is the mean of the three results and sits inside the spread of the data, so the evidence **supports** the hypothesis. It does not **prove** it: three concordant results are consistent with $0.100\ \text{mol L}^{-1}$, but the true value is only ever bounded, never confirmed absolutely. One **assumption** is that the indicator's colour change coincides with the equivalence point; a **limitation** is that the conclusion applies only to this solution under these conditions. The **logbook** — a dated, contemporaneous record of every raw burette reading, signed as the student's own work — authenticates the data: it shows the results are genuinely those generated in the investigation and lets another chemist trace or reproduce them.

$\therefore$ the evidence supports (but does not prove) the hypothesis that the concentration is $0.100\ \text{mol L}^{-1}$.

Practice questions

Question 1 — In a titration, a student records five titres, in mL: 22.30, 21.65, 21.75, 21.70, 21.70.

a. Identify the outlier and justify excluding it.
b. Calculate the mean titre that should be used.
c. Explain why averaging the concordant titres, rather than all five, improves the accuracy of the measurement result.

Question 2 — Classify each of the following as a **random** or a **systematic** error, and state its effect on the results.

a. In each burette reading a student must estimate the last figure between two graduations, and that estimate varies slightly from one reading to the next.
b. A balance is not zeroed and reads $0.02\ \text{g}$ too high for every mass.
c. A burette is rinsed with distilled water (not the titrant) before filling, slightly diluting the solution it delivers.

Question 3 —

a. A $50\ \text{mL}$ burette is graduated every $0.1\ \text{mL}$. State its resolution and the uncertainty in a single volume reading.

b. A digital balance displays a mass of $2.47\,\text{g}$. State its resolution and express the mass with its uncertainty.
c. A digital thermometer reading to $0.1\,^\circ\text{C}$ is used to measure the same temperature three times: 24.6, 24.8 and $24.7\,^\circ\text{C}$. State the measurement result and comment on what the spread of the readings shows about the precision.

Question 4 — Two students each determine the same concentration three times. The accepted value is $0.150\,\text{mol L}^{-1}$.

Student A: 0.168, 0.169, $0.170\,\text{mol L}^{-1}$. Student B: 0.145, 0.154, $0.151\,\text{mol L}^{-1}$.

a. Whose results are more precise?
b. Whose results are more accurate?
c. What type of error is present in Student A's results, and how could it be reduced?

Question 5 — In a colorimetric investigation the absorbance of five standard solutions is measured:

concentration/ mmol L^{-1}	1.0	2.0	3.0	4.0	5.0
absorbance	0.15	0.25	0.30	0.45	0.55

a. Identify the outlier.
b. Ignoring the outlier, determine the gradient and the vertical intercept of the line of best fit, and state what the gradient represents.
c. Use the line to find the concentration of a sample of absorbance 0.40, and state one limitation of this determination.

Question 6 —

a. Explain the difference between saying evidence "supports" a hypothesis and saying it "proves" a hypothesis.
b. State the purpose of a logbook in authenticating primary data.
c. A student concludes "the rate of reaction increases with temperature" from data collected between $20\,^\circ\text{C}$ and $50\,^\circ\text{C}$. Identify one limitation that bounds this conclusion.

Question 7 — Three repeated weighings of a sample on a balance reading to $0.01\,\text{g}$ give 1.02, 1.03 and $1.04\,\text{g}$. State the measurement result and explain what the spread of the three weighings indicates about the precision of the weighing and the type of error present.

Question 8 — A student records five titres, in mL: 27.20, 26.15, 26.05, 26.10, 26.10. Calculate the mean titre that should be reported, and justify excluding any value.

Question 9 — Classify each error as random or systematic, and state which of them is reduced by repeating the measurement: (i) a stopwatch is started a little late by differing amounts each trial; (ii) a thermometer reads $1.0\,^\circ\text{C}$ below the true temperature at every reading; (iii) a student reads every burette meniscus from slightly below the level of the scale.

Question 10 — A digital pH meter displays 4.36, and a measuring cylinder graduated every $1\,\text{mL}$ reads $25\,\text{mL}$. Express each measurement with its uncertainty.

Question 11 — A student's three repeated results agree closely with one another but differ noticeably from the accepted value. Comment on the precision and the accuracy of the results, and identify the most likely type of error.

Question 12 — In an investigation of reaction rate, a student measures the volume of carbon dioxide produced as marble chips (calcium carbonate) react with excess dilute hydrochloric acid. The volume is plotted against time and, over the first $30\,\text{s}$, a line of best fit is drawn with a gradient of $3.0\,\text{mL s}^{-1}$. Write a balanced equation, with states, for the reaction; state what the gradient represents; and explain why a line of best fit is drawn rather than joining each point to the next.

Question 13 — A student writes: “My single experiment proves that the reaction is exothermic.” Evaluate this statement with reference to the nature of scientific evidence.

Question 14 — Describe two features of a properly kept logbook that allow it to authenticate the primary data generated in an investigation.

Question 15 — In a titration carried out to find the concentration of an acid, state one assumption made in the analysis and one limitation that bounds the conclusion that can be drawn.

Question 16 — A hypothesis predicts a molar enthalpy of solution of $25.0\,\text{kJ mol}^{-1}$. Three trials give 24.1, 24.6 and $25.1\,\text{kJ mol}^{-1}$. Determine the measurement result and state, with reference to the spread of the trials, whether the evidence supports the prediction.

Question 17 — Five titres, in mL, are recorded: 19.95, 20.05, 20.60, 20.00, 20.00. The accepted titre for the solution is $20.15\,\text{mL}$. Calculate the mean titre (justifying any exclusion), and state — with reference to the spread of the concordant titres — whether the evidence supports the accepted value.

Question 18 — A student tests the prediction that a standardised solution will require a titre of $23.40\,\text{mL}$. Four titres, in mL, are recorded: 23.45, 23.50, 23.55, 24.20.

a. Calculate the mean titre, justifying any value you exclude.
b. Comment on the precision of the concordant titres, referring to their spread.
c. Does the evidence support the prediction? Refer to the spread of the concordant titres in your answer.
d. Explain how the student’s logbook authenticates these titres, and state one way to improve the precision of the result.

Answers

1 a 22.30 (not concordant with the group, about $0.6\,\text{mL}$ high); b $21.70\,\text{mL}$; c excluding the anomalous result removes a large one-off (random) deviation, giving a mean closer to the true value — a more accurate measurement result. **2** a random — lowers precision, reduced by repeats; b systematic — lowers accuracy, not reduced by repeats; c systematic — every titre reads too large, lowering accuracy. **3** a resolution $0.1\,\text{mL}$, $\pm 0.05\,\text{mL}$; b resolution $0.01\,\text{g}$, $(2.47 \pm 0.01)\,\text{g}$; c measurement result $24.7\,°\text{C}$; the readings span only $0.2\,°\text{C}$ (two display steps), so they are closely grouped — precise. **4** a Student A; b Student B; c a systematic error; identify and remove its source (e.g. recalibrate the balance / restandardise the solution). **5** a $(3.0, 0.30)$; b gradient $0.10\,\text{L mmol}^{-1}$ (absorbance per mmol L^{-1}), intercept 0.05; the gradient is the absorbance produced per unit concentration; c $3.5\,\text{mmol L}^{-1}$; valid only within the calibrated range (assumes the linear relationship continues). **6** a "supports" = consistent with, given the spread of the data; proof is impossible, so conclusions stay provisional; b a dated, contemporaneous, authenticated record of the raw data; c the conclusion holds only over 20–50 °C and cannot be extrapolated beyond that range. **7** measurement result $1.03\,\text{g}$; the weighings span $0.02\,\text{g}$ (two steps of the balance's last digit), so they are closely grouped and precise; the small variation is random error. **8** $26.10\,\text{mL}$ (27.20 excluded — not concordant). **9** (i) random (reduced by repeats); (ii) systematic (not reduced by repeats); (iii) systematic — parallax (not reduced by repeats). Only (i) is reduced by repeating. **10** (4.36 ± 0.01); $(25.0 \pm 0.5)\,\text{mL}$. **11** precise but not accurate; a systematic error. **12** $\text{CaCO}_3(\text{s}) + 2\,\text{HCl}(\text{aq}) \rightarrow \text{CaCl}_2(\text{aq}) + \text{H}_2\text{O}(\text{l}) + \text{CO}_2(\text{g})$; the gradient is the rate of gas production, $3.0\,\text{mL s}^{-1}$; a line of best fit averages out the random scatter of individual points. **13** a single trial cannot prove anything; evidence can support but never prove, and one result cannot be shown to be repeatable — the claim is provisional at best. **14** e.g. dated, contemporaneous entries of the raw readings, and a signature (and record of any method changes) — see solutions. **15** assumption e.g. the endpoint coincides with the equivalence point; limitation e.g. the result applies only to this solution under these conditions. **16** measurement result $24.6\,\text{kJ mol}^{-1}$; the trials are widely spread (24.1–25.1, i.e. $1.0\,\text{kJ mol}^{-1}$) and 25.0 lies inside that spread, so the evidence supports the prediction — but only weakly, because the trials are imprecise. **17** $20.00\,\text{mL}$ (20.60 excluded); the concordant titres span only 19.95–$20.05\,\text{mL}$ and every one is below $20.15\,\text{mL}$, so the evidence does not support the accepted value (a systematic error is likely). **18** a $23.50\,\text{mL}$ (24.20 excluded); b spread of only $0.10\,\text{mL}$ — concordant, so precise; c no — all three concordant titres lie above $23.40\,\text{mL}$, which sits outside their spread; d dated authenticated record of raw readings; improve precision by taking more repeat titrations.

Fully-worked solutions

Question 1 a. The concordant titres cluster at 21.65–$21.75\,\text{mL}$ (agreeing within $0.10\,\text{mL}$). The value $22.30\,\text{mL}$ lies about $0.6\,\text{mL}$ above this group, so it is an outlier and is excluded — most likely the end point was overshot. b. $\text{mean} = \dfrac{21.65 + 21.75 + 21.70 + 21.70}{4} = \dfrac{86.80}{4} = 21.70\,\text{mL}$. c. Averaging all five would include the anomalous $22.30\,\text{mL}$, biasing the mean upward ($109.10/5 = 21.82\,\text{mL}$). Averaging only the concordant titres removes that one-off deviation and gives a mean that lies closer to the true titre — that is, a more **accurate** measurement result.

Question 2 a. **Random** error: the last figure of a burette reading has to be estimated between two graduations, and that estimate falls sometimes above and sometimes below the true value in an unpredictable way. It lowers the precision (increases scatter) and its effect is reduced by taking repeats and averaging. b. **Systematic** error: the balance is biased $+0.02\,\text{g}$ on every reading, so all masses are too high by the same amount. It lowers the accuracy and is not reduced by repeating — every repeat carries the same bias. It is corrected by zeroing (calibrating) the balance. c. **Systematic** error: rinsing with water dilutes the titrant, so a slightly larger volume is needed each time and every titre is too large in the same direction. This lowers accuracy; rinsing the burette with the titrant removes the error.

Question 3 a. Resolution = $0.1\,\text{mL}$ (the smallest graduation). A single reading is estimated to ± half a division, so the uncertainty is $\pm 0.05\,\text{mL}$. b. Resolution = $0.01\,\text{g}$ (the last displayed digit). The mass is quoted to ± the last digit: $(2.47 \pm 0.01)\,\text{g}$. c. The measurement result is the mean of the repeats: $\dfrac{24.6 + 24.8 + 24.7}{3} = 24.7\,°\text{C}$. The three readings span only 24.6–$24.8\,°\text{C}$ — two steps of the $0.1\,°\text{C}$ display — so they agree closely with one another and the measurement is **precise**. (Precision says nothing about accuracy: a mis-calibrated thermometer could give three closely grouped readings that are all wrong.)

Question 4 a. **Student A**. A's results span only 0.168–0.170 (range $0.002\,\text{mol L}^{-1}$), much tighter than B's 0.145–0.154 (range $0.009\,\text{mol L}^{-1}$), so A's data is more precise. b. **Student B**. B's mean is $0.145+0.154+0.151/3=0.150\,\text{mol L}^{-1}$, equal to the accepted value; A's mean is $0.168+0.169+0.170/3=0.169\,\text{mol L}^{-1}$, about $0.019\,\text{mol L}^{-1}$ too high. B's results are more accurate. c. A's results are precise but consistently high, which is the signature of a **systematic** error. Repeating will not help; the source (e.g. an un-standardised solution or an un-zeroed balance) must be identified and corrected.

Question 5 a. The point $(3.0, 0.30)$ lies below the straight-line trend set by the other four points, so it is the outlier. b. Using the four points that lie on the trend, $(1.0, 0.15), (2.0, 0.25), (4.0, 0.45), (5.0, 0.55)$, the line of best fit is $A = 0.10c + 0.05$: the **gradient** is $0.10\,\text{L mmol}^{-1}$ (i.e. the absorbance rises by 0.10 for each $1\,\text{mmol L}^{-1}$ increase in concentration) and the **intercept** is 0.05 (the small absorbance of the blank). The gradient represents the absorbance produced per unit concentration. c. Rearranging, $c = \dfrac{A-0.05}{0.10} = \dfrac{0.40-0.05}{0.10} = 3.5\,\text{mmol L}^{-1}$. This is valid only if the sample lies within the calibrated range 1–$5\,\text{mmol L}^{-1}$; the determination assumes the linear relationship between absorbance and concentration continues to hold, and it would not be justified if extrapolated beyond that range.

Question 6 a. To say the evidence **supports** a hypothesis means the results are consistent with it, given the spread of the measurements. To say it **proves** a hypothesis claims certainty — but no finite set of results can do this, because a future measurement could always disagree. Scientific conclusions are therefore provisional: evidence can support or refute, but never prove. b. A logbook is a dated, contemporaneous record of the raw data and observations, kept as the work proceeds and signed as the student's own. It authenticates the primary data — showing it is genuinely what was generated, not altered or fabricated — and allows the data to be checked or reproduced. c. The data was collected only between $20\,°\text{C}$ and $50\,°\text{C}$, so the conclusion is bounded to that range. It cannot be assumed to hold outside it (for example, at high temperatures a catalyst or reactant might behave differently), so extrapolating beyond $50\,°\text{C}$ is not justified by this evidence.

Question 7 The measurement result is the mean of the repeats: $\dfrac{1.02+1.03+1.04}{3} = \dfrac{3.09}{3} = 1.03\,\text{g}$. The three weighings span 1.02–$1.04\,\text{g}$, i.e. two steps of the balance's last displayed digit, so they agree closely with one another and the weighing is **precise**. The small trial-to-trial variation, falling both above and below the mean, is a **random** error; averaging the repeats reduces its effect on the measurement result.

Question 8 The titres 26.05–$26.15\,\text{mL}$ are concordant (within $0.10\,\text{mL}$); $27.20\,\text{mL}$ lies about $1.1\,\text{mL}$ above them and is an outlier, so it is excluded. $\text{mean} = \dfrac{26.15+26.05+26.10+26.10}{4} = \dfrac{104.40}{4} = 26.10\,\text{mL}$.

Question 9 (i) The lateness varies unpredictably from trial to trial, so it is a **random** error; its effect is reduced by repeating the measurement and averaging. (ii) The thermometer is $1.0\,°\text{C}$ low on every reading — a constant bias in one direction — so it is a **systematic** error; repeating does not reduce it, because every reading carries the same offset. (iii) Reading the meniscus from below the level of the scale is **parallax**, an inappropriate method that biases every reading in the same direction, so it is a **systematic** error; it is removed by reading the scale at eye level, not by repeating. Of the three, only the random error (i) is reduced by repeating and averaging.

Question 10 The pH meter is digital, so its resolution is the last digit, 0.01: the reading is (4.36 ± 0.01). The measuring cylinder is analogue with $1\,\text{mL}$ graduations, so its reading uncertainty is ± half a division = $\pm 0.5\,\text{mL}$: the volume is $(25.0 \pm 0.5)\,\text{mL}$.

Question 11 Because the repeated results agree closely with one another, the data is **precise**. Because they differ noticeably from the accepted value, the data is **not accurate**. Precise-but-inaccurate results are the signature of a **systematic** error, which shifts every result in the same direction and is not removed by repeating.

Question 12

$$\text{CaCO}_3(\text{s}) + 2\,\text{HCl}(\text{aq}) \rightarrow \text{CaCl}_2(\text{aq}) + \text{H}_2\text{O}(\text{l}) + \text{CO}_2(\text{g})$$

The gradient of a volume-against-time graph is the change in gas volume per unit time — that is, the **rate of carbon dioxide production**, here $3.0\,\text{mL s}^{-1}$. A line of best fit is drawn rather than joining consecutive points because each measured point carries random scatter; the best-fit line balances the points and averages out that scatter, giving a better estimate of the underlying relationship (and its gradient) than a dot-to-dot join, which would treat the noise as if it were real.

Question 13 The statement is not justified. First, a single trial cannot be shown to be **repeatable**: without repeats there is no way to judge the precision of the result or to detect an anomaly. Second, and more fundamentally, evidence can **support** or **refute** a hypothesis but can never **prove** it — a scientific conclusion is always provisional and open to revision by further evidence. A defensible claim would be that the results **support** the hypothesis that the reaction is exothermic, pending repetition of the measurement and, ideally, reproduction of the result by another analyst.

Question 14 Any two of: (1) entries are **dated** and made **contemporaneously** — recorded at the time the data is generated, not written up later; (2) the **raw readings** (e.g. every burette reading) are recorded directly, not just processed results, so they can be re-checked; (3) the logbook is **signed** (and where required witnessed) as the student's own work; (4) any **changes or adaptations** to the method are noted as they happen. Together these establish that the data is genuine, traceable and reproducible.

Question 15 Assumption (any one): the indicator's colour change marks the equivalence point; the standard solution's concentration is exactly known; the pipette and burette deliver their stated volumes. Limitation (any one): the result applies only to this particular solution under the conditions tested; the precision is limited by the resolution of the glassware; a systematic error in the standard would bias the result. These bound the conclusion to what the data can actually support.

Question 16 $\text{measurement result} = \text{mean} = \dfrac{24.1 + 24.6 + 25.1}{3} = \dfrac{73.8}{3} = 24.6\,\text{kJ mol}^{-1}$. The three trials span 24.1–$25.1\,\text{kJ mol}^{-1}$ — a spread of $1.0\,\text{kJ mol}^{-1}$, about $4\,\%$ of the mean — so they are not closely grouped and the data is **imprecise**. The predicted value $25.0\,\text{kJ mol}^{-1}$ lies inside that spread and is close to the mean, so the evidence is **consistent with** (supports) the prediction. The support is weak, however: because the trials scatter so widely, further repeats would be needed before the prediction could be confidently accepted.

Question 17 The titres 19.95–$20.05\,\text{mL}$ are concordant; $20.60\,\text{mL}$ is an outlier and is excluded. $\text{mean} = \dfrac{19.95 + 20.05 + 20.00 + 20.00}{4} = \dfrac{80.00}{4} = 20.00\,\text{mL}$. The concordant titres span only 19.95–$20.05\,\text{mL}$, so the titrations are **precise**; but every one of them is smaller than the accepted titre $20.15\,\text{mL}$, which lies clearly outside that spread. The evidence therefore does **not** support the accepted value: precise results that are consistently low in the same direction are the signature of a systematic error (for example, an under-standardised titrant), which repeating would not remove.

Question 18 a. The titres 23.45, 23.50 and $23.55\,\text{mL}$ are concordant (within $0.10\,\text{mL}$); $24.20\,\text{mL}$ lies about $0.7\,\text{mL}$ higher and is an outlier, so it is excluded. $\text{mean} = \dfrac{23.45 + 23.50 + 23.55}{3} = \dfrac{70.50}{3} = 23.50\,\text{mL}$. b. The three concordant titres span $23.55 - 23.45 = 0.10\,\text{mL}$, the usual concordance window for titres. They agree closely with one another, so the titrations are **precise** (the fourth titre, $24.20\,\text{mL}$, is a one-off outlier rather than evidence of poor precision). c. All three concordant titres are larger than the predicted $23.40\,\text{mL}$, which lies outside the spread 23.45–$23.55\,\text{mL}$ of the measured values, so the evidence does **not** support the prediction. The discrepancy is small — only about the width of the concordance window — so before ruling the prediction out the student should carry out more repeats and have the titration reproduced independently, and should check for a systematic error such as consistently over-running the end point. d. The logbook records every raw burette reading, dated and entered at the time of the titration and signed as the student's own work, so it authenticates the titres as genuinely generated in the investigation and lets another analyst trace or reproduce them. Precision could be improved by carrying out more repeat titrations, reducing the effect of random error on the mean.

Chapter 13 Scientific investigations and data analysis — 13C Communicating chemistry

Theory

An investigation is only complete once its findings are **communicated**. In chemistry this is done with a shared set of conventions — significant figures, units, symbols, tables and graphs — so that any reader can judge, reproduce and build on the work. The Unit 4 investigation is communicated as a **scientific poster**, and its findings are expressed as an **evidence-based conclusion**. This subtopic covers the conventions of scientific communication and the discipline of drawing a conclusion that says no more, and no less, than the data allow.

Significant figures. The **significant figures** of a measurement are the digits that carry real information about how precisely the quantity is known. The rules are:

- all **non-zero** digits are significant;
- zeros **between** non-zero digits (captive zeros) are significant;
- **leading** zeros (before the first non-zero digit) are *never* significant — they only fix the position of the decimal point, so 0.0042 has 2 significant figures;
- **trailing** zeros are significant **only if the number contains a decimal point** — 25.0 has 3 significant figures, but a plain 250 is ambiguous (write 2.5×10^2 for 2 figures or 2.50×10^2 for 3);
- exact or counted values (for example a mole ratio, or "3 trials") and defined constants are treated as having unlimited significant figures and never limit an answer.

When quantities are combined, the answer must not imply more precision than the data provide:

- in **multiplication and division**, the answer carries the **same number of significant figures** as the datum with the **fewest** significant figures;
- in **addition and subtraction**, the answer is rounded to the **fewest number of decimal places** among the data;
- round only at the **final** step, keeping extra figures in intermediate working.

The key idea is that an answer's number of significant figures **reflects the precision of the data** — calculation cannot create precision that the measurements do not have. Reporting a titre-based concentration as $0.083\,642\ \text{mol L}^{-1}$ claims **five** significant figures (the leading zeros do not count, so the significant digits are $8, 3, 6, 4, 2$), yet titres read to two decimal places and a standard solution known to three figures justify at most **three**.

SI units, prefixes and units on every quantity. Chemistry uses the **International System of Units (SI)**. The base units include the metre (m), kilogram (kg), second (s), mole (mol), kelvin (K) and ampere (A); derived units such as the joule (J), volt (V), pascal (Pa) and coulomb (C) are built from these. **Metric prefixes** scale a unit by a power of ten:

Prefix	Symbol	Factor	Prefix	Symbol	Factor
giga	G	10^{9}	centi	c	10^{-2}
mega	M	10^{6}	milli	m	10^{-3}
kilo	k	10^{3}	micro	μ	10^{-6}
deci	d	10^{-1}	nano	n	10^{-9}
			pico	p	10^{-12}

Every measured or calculated quantity must be reported **with its unit**: a bare number is meaningless and, in an examination, loses the mark. Changing the prefix changes only the number's magnitude, never its number of significant figures ($4.2 \times 10^{-5}\ \text{g L}^{-1} = 42\ \mu\text{g L}^{-1}$, still 2 significant figures).

Symbols, formulae, tables and graphs. Scientific communication is precise about representation. Chemical symbols are case-sensitive (Co is cobalt; CO is carbon monoxide); formulae use correct subscripts and charges (SO_4^{2-}); equations are **balanced and carry state symbols**; organic compounds are named by **IUPAC** rules. Quantity symbols

are written in italics and unit symbols upright. In a **table**, each column is headed by the *physical quantity and its unit* (for example "Temperature / °C"), so that the entries are pure numbers quoted to a consistent number of decimal places. In a **graph**, both axes are labelled with the quantity and unit, a sensible scale is chosen, and a line or curve of best fit (or a calibration curve) is drawn. These conventions let a reader interpret the data unambiguously.

Evidence, inference and opinion. A scientific argument is built only from **evidence** — the data and observations that were generated or collated. An **inference** is a reasoned *interpretation* of evidence (for example, using collision theory to explain a trend); it is legitimate in a discussion provided it is clearly tied to the evidence. An **opinion** (or anecdote) is a personal value judgement that no data support ("this was the most interesting result"). Only evidence and sound inference can support or refute a hypothesis; opinions and anecdotes cannot. A central skill is to read a statement and decide which of these it is.

Drawing conclusions and the discussion. The **discussion** interprets the results: it analyses patterns, evaluates sources of random and systematic error, states the assumptions and limitations of the method, considers validity, and compares the findings with chemical theory. The **conclusion** is then a concise statement that:

- directly **answers the research question** and relates to the **aim**;
- is **bounded by the data** — it claims no more than the evidence shows, and does not extend to conditions, quantities or reactions that were not investigated;
- states the **extent to which the hypothesis is supported or refuted**, acknowledging the uncertainty in the data.

A conclusion that over-generalises ("this is true for all reactions"), introduces an unmeasured claim ("therefore it is the healthiest"), or merely restates prior knowledge without using the data, is not evidence-based. Words such as *prove* are avoided; evidence *supports* or *is consistent with* a hypothesis to a stated degree.

The scientific poster. VCAA requires the Unit 4 investigation to be presented as a **scientific poster**, produced electronically or in hard copy. The poster **must not exceed 600 words**; supporting text in tables, graphs, image captions, references and acknowledgements is not counted. Posters are **not limited to a particular paper size or number of panels** — the priorities are conciseness, clarity and legibility. VCAA prescribes the format: **title** (the question under investigation, with the student's name); **introduction** (the reason for undertaking the investigation, a clear aim, the hypothesis or prediction, and the relevant background chemistry); **methodology and methods** (the methodology chosen to address the question, and a summary of the data-generation and data-analysis methods, including the independent, dependent and controlled variables); **results** (the generated data presented so that trends, patterns and relationships are clear); **discussion** (interpretation and evaluation of the analysed data, limitations of the data and methods with suggested improvements, cross-referencing of the results to chemical concepts, and whether the findings support the hypothesis); **conclusion** (a response to the investigation question and the extent to which it has been answered, **with no new information introduced**); and **references and acknowledgements**. At the **centre** of the poster, occupying **20–25%** of the poster space, sits the **communication statement**: a **one-sentence summary of the major finding** that answers the investigation question. It is not an abstract — it reports the finding, and the reader who reads nothing else still learns what the investigation found. Alongside the poster, the **logbook** — recording the investigation planning, the identification and management of risks, all raw data, and the treatment of any outliers — is also assessed.

Referencing, acknowledgement and ethical reporting. Any value, idea, image or method taken from another source must be **acknowledged**: cited in-text at the point of use and listed in full under **References** in the poster's *References and acknowledgements* section (author or organisation, year, title, source, and for online sources the date accessed), following a standard referencing style. Reproducing a diagram or passage as one's own is plagiarism. Assistance from a teacher, technician or peer is recognised under **Acknowledgements** in the same section. Reporting must be **ethical and authentic**: all raw data are recorded in a **logbook** (which authenticates the work), and every result is reported honestly. Data are never fabricated, and a result may be excluded only if there is a justified reason to treat it as an **outlier** — the exclusion must be *stated and justified*, with the raw value remaining in the logbook. Selectively deleting data to make a result look "nicer" is falsification and is scientifically dishonest.

Worked examples

Worked Example 1 — In a calorimetry experiment a student measures the mass of water as $50.0\,\text{g}$, uses $c = 4.18\,\text{J g}^{-1}\,^{\circ}\text{C}^{-1}$, and records a temperature rise of $4.2\,^{\circ}\text{C}$. Calculate the heat absorbed, $q = m\,c\,\Delta T$, and express it to the correct number of significant figures, explaining your choice.

Working	Comment
$m = 50.0\,\text{g}$ (3 sf), $c = 4.18\,\text{J g}^{-1}\,^{\circ}\text{C}^{-1}$ (3 sf), $\Delta T = 4.2\,^{\circ}\text{C}$ (2 sf)	List each datum with the significant figures it carries.
$q = m\,c\,\Delta T = 50.0 \times 4.18 \times 4.2 = 877.8\,\text{J}$	Full-precision product; do not round yet.
Fewest significant figures among the data = 2 (from ΔT)	In multiplication the least-precise datum limits the answer.
$q = 8.8 \times 10^{2}\,\text{J} = 0.88\,\text{kJ}$ (2 sf)	Standard form makes the two significant figures unambiguous; the answer can be no more precise than ΔT.

Worked Example 2 — A water analyst reports a lead(II) concentration of $0.000\,042\,\text{g L}^{-1}$. Express this concentration (a) in mg L^{-1} using the prefix *milli*, (b) in $\mu\text{g L}^{-1}$ using the prefix *micro*, and (c) in standard form in g L^{-1}, each to the correct number of significant figures.

Working	Comment
$0.000\,042\,\text{g L}^{-1} = 4.2 \times 10^{-5}\,\text{g L}^{-1}$ (2 sf)	Standard form first; the value carries 2 significant figures.
milli = 10^{-3}: $4.2 \times 10^{-5}\,\text{g L}^{-1} = 4.2 \times 10^{-2}\,\text{mg L}^{-1} = 0.042\,\text{mg L}^{-1}$	Multiply the number by 10^{3} to change g → mg.
micro = 10^{-6}: $4.2 \times 10^{-5}\,\text{g L}^{-1} = 42\,\mu\text{g L}^{-1}$	Multiply the number by 10^{6} to change g → μg.
All three forms keep 2 significant figures	Converting units changes the magnitude, never the number of significant figures.

Worked Example 3 — A student investigated the aim *"To determine whether increasing temperature increases the rate of the reaction between sodium thiosulfate and hydrochloric acid,"* timing how long a printed cross took to disappear. The mean results were: 20 °C → 62 s; 40 °C → 28 s; 60 °C → 13 s. Classify each statement below as **evidence**, an **inference**, or an **opinion**, then write an evidence-based conclusion.

(i) "As the temperature rose from 20 °C to 60 °C, the mean time for the cross to disappear fell from 62 s to 13 s."
(ii) "The higher temperature gives the particles more kinetic energy, so more collisions exceed the activation energy and the rate increases."
(iii) "The 60 °C trial was the most satisfying to watch."

Statement (i) reports the measured data, so it is **evidence**. Statement (ii) interprets that evidence using collision theory rather than reporting a measurement, so it is an **inference**. Statement (iii) is a personal value judgement that no data support, so it is an **opinion**. A shorter time means a faster rate, so the rate increased with temperature across the range tested.

∴ the data support the hypothesis that increasing temperature increases the reaction rate: raising the temperature from 20 °C to 60 °C reduced the mean reaction time from 62 s to 13 s. The conclusion is bounded by the conditions investigated — this reaction over the 20–60 °C range — and cannot be extended to other reactions or to temperatures outside this range.

Worked Example 4 — A group is assembling a scientific poster for a Unit 4 investigation into the concentration of ethanoic acid in vinegar by titration. **(a)** State the poster section in which each item belongs: (i) the aim and hypothesis; (ii) the average titre and the titration table; (iii) an evaluation of random and systematic errors; (iv) a one-sentence statement of the extent to which the hypothesis was supported; (v) the data source used for a literature value.

(b) The Results section reports the concentration as "$0.08236\,\text{M}$". The three concordant titres were each recorded to two decimal places (e.g. $15.20\,\text{mL}$). Comment on the reporting.

(a)

(i) **Introduction**; (ii) **Results**; (iii) **Discussion**; (iv) **Conclusion**; (v) **References and acknowledgements** — VCAA combines the two, so a cited data source and the recognition of personal assistance both appear in this section.

(b) The **unit is acceptable**: "M" is standard VCAA notation for mol L^{-1} — the VCE Chemistry Data Book uses it ("the concentration of all soluble species is assumed to be $1.00\,\text{M}$") and so do the examination papers — although writing mol L^{-1} in full is equally correct. What is wrong is the **precision**: the titres carry four significant figures and the standard concentration and pipette volume typically three, so the final concentration is justified to at most three significant figures, yet 0.08236 quotes four (the leading zeros are not significant).

$\therefore$ each quantity on a poster must carry its unit and be quoted to a number of significant figures justified by the data — here the concentration should read $8.24 \times 10^{-2}\,\text{mol L}^{-1}$ (equivalently $0.0824\,\text{M}$), to 3 sf.

Practice questions

Question 1 — During an experiment a student records the mass of a sample as $3.450\,\text{g}$, its volume as $12.0\,\text{mL}$, and a thermometer reading of $25.5\,°\text{C}$.

a. State the number of significant figures in each of the three measurements.
b. The density is mass ÷ volume. Calculate the density and express it to the correct number of significant figures.
c. State which measurement limits the precision of the density, and why.

Question 2 — Convert each quantity, expressing the answer with the prefix or unit indicated.

a. A concentration of $0.0056\,\text{mol L}^{-1}$ into mmol L^{-1}.
b. A mass of $2500\,\text{mg}$ into g, and then into kg (in standard form).
c. A volume of $750\,\text{mL}$ into L, and a wavelength of $0.000\,000\,540\,\text{m}$ into nm.

Question 3 — During an investigation into the effect of catalyst surface area on the decomposition of hydrogen peroxide, a student wrote the following. Classify each statement as **evidence**, an **inference**, or an **opinion**.

a. "Using powdered manganese dioxide, 48 mL of oxygen was collected in 30 s; using a single lump of the same mass, 21 mL was collected in 30 s."
b. "The powder produced more gas because its larger surface area provides more sites for the reaction."
c. "Powdered catalysts are the best choice for any experiment." Justify your classification.

Question 4 — An investigation had the aim *"To determine whether the citric acid content of three brands of lemon drink differs."* Titration gave: Brand A $0.82\,\text{g}$, Brand B $0.79\,\text{g}$, Brand C $0.55\,\text{g}$ of citric acid per $100\,\text{mL}$ (each $\pm 0.02\,\text{g}$). A student concluded: *"Brand C is the healthiest lemon drink because it contains the least acid."*

a. State, with a reason, whether this conclusion is bounded by the data.
b. State whether the conclusion addresses the aim.
c. Write an improved conclusion that is bounded by the data and addresses the aim.

Question 5 — A scientific poster reports a Unit 4 investigation. State the poster section in which each of the following belongs.

a. The research question and the hypothesis.
b. A calibration curve of absorbance versus concentration, and a table of raw titres.
c. An evaluation of the random and systematic errors and how they affected the results.

Question 6 — A student is preparing the poster.

a. The student uses a literature melting-point value from a chemistry data source. State what must be done so the source is properly acknowledged.

b. A laboratory technician prepared the standard solutions used. State how this should be recognised on the poster.
c. One titre was clearly anomalous. Explain what the student should do with this result in order to report ethically and authentically.

Question 7 — A gas sample occupies $0.250\ \text{L}$ and contains $0.0104\ \text{mol}$. Its molar volume is volume ÷ amount. Calculate the molar volume, giving your answer to the correct number of significant figures.

Question 8 — A burette reads $23.40\ \text{mL}$ at the start of a titration and $47.6\ \text{mL}$ at the end (the final reading was estimated to one decimal place). Calculate the titre, express it to the correct number of significant figures, and justify your choice.

Question 9 — An analyst reports the concentration of mercury in a water sample as $3.5 \times 10^{-6}\ \text{g L}^{-1}$. Express this concentration in (i) $\mu\text{g L}^{-1}$ and (ii) ng L^{-1}.

Question 10 — The following table appears in the Results section of a poster. State three faults in the way the data are presented, and give the corrected form.

Temp	Volume of gas
20	12
30	25.0
40	48

Question 11 — Classify each of the following statements from a report as **evidence**, an **inference**, or an **opinion**.

(i) "The solution changed from colourless to pink at a mean titre of $18.65\ \text{mL}$."
(ii) "The pink colour shows the end point had been reached, so the acid had been neutralised."
(iii) "This was the easiest titration to perform."

Question 12 — An investigation tested the hypothesis that adding a catalyst increases the rate of a reaction. The mean reaction time fell from $85\ \text{s}$ (no catalyst) to $32\ \text{s}$ (with catalyst). A student wrote: *"The catalyst made the reaction go heaps faster and is definitely the best way to speed up any reaction."* Rewrite this conclusion so that it is evidence-based, quantitative and bounded by the data.

Question 13 — An investigation had the aim *"To determine the concentration of ethanoic acid in a sample of vinegar."* The titration gave a mean concentration of $0.82\ \text{mol L}^{-1}$ ($\pm 0.03\ \text{mol L}^{-1}$). A student's conclusion read: *"Vinegar is an acid and should be handled with care."* Evaluate whether this conclusion addresses the aim and is supported by the data, then write a conclusion that does.

Question 14 — A group is finalising the poster for its Unit 4 investigation.

a. The draft contains 940 words in total, of which 180 words are in table entries, figure captions, references and acknowledgements. State, with reasoning, whether the poster meets the VCAA length requirement.
b. The group intends to print the poster on a single A4 sheet "because VCAA requires a one-page poster". Evaluate this claim.
c. Name the poster section in which the independent, dependent and controlled variables and the procedure belong.
d. The group wants to place a comparison of its result with a published literature value in the Conclusion, to show that its result is reasonable. State why this is not appropriate there, and name the section in which the comparison belongs.

Question 15 — a. State the form the communication statement on a VCAA scientific poster must take, where it is placed, and roughly what proportion of the poster it occupies. b. State what a reader who reads nothing else on the poster should learn from it. c. For the investigation below, write a suitable communication statement. *Investigation: the vitamin C content of fresh and cooked broccoli was compared by redox titration; fresh broccoli contained* $62\ \text{mg}$ *per* $100\ \text{g}$ *and cooked broccoli* $38\ \text{mg}$ *per* $100\ \text{g}$*; the hypothesis was that cooking lowers vitamin C content.*

Question 16 — A student quotes a literature value for the density of ethanol taken from a chemistry data source, and reproduces a diagram of an apparatus from a website. Explain, with reference to scientific referencing conventions, what the student must do in each case.

Question 17 — While analysing data, a student notices that including all five trials gives a mean that does not fit the expected trend, but deleting the two lowest results produces a "nicer" graph. Explain why deleting the data in this way breaches ethical and authentic reporting, and describe the correct procedure.

Question 18 — A poster's Discussion and Conclusion for an investigation into whether concentration affects reaction rate reads: *"At* $0.10\,\text{mol L}^{-1}$ *the rate was 0.0021 and at* $0.20\,\text{mol L}^{-1}$ *it was 0.0043. Concentration always doubles the rate, which proves my hypothesis and shows this is true for all reactions."* Identify three faults in the communication and conclusion conventions, and write a corrected conclusion.

Answers

1 a $3.450 \rightarrow 4$; $12.0 \rightarrow 3$; $25.5 \rightarrow 3$. b $0.288\ \text{g mL}^{-1}$. c the volume $12.0\ \text{mL}$ (3 sf — the fewest). **2** a $5.6\ \text{mmol L}^{-1}$. b $2.5\ \text{g} = 2.5 \times 10^{-3}\ \text{kg}$. c $0.750\ \text{L}$; $540\ \text{nm}$. **3** a evidence; b inference; c opinion (an unsupported over-generalisation). **4** a No — "healthiest" is not measured and adds an unsupported judgement. b No — it goes beyond the aim (acid content). c see solution. **5** a Introduction; b Results; c Discussion. **6** a in-text citation at point of use + a full entry in References; b in the Acknowledgements; c record it in the logbook and exclude it only as a justified outlier, stated openly — never delete or alter it silently. **7** $24.0\ \text{L mol}^{-1}$ (3 sf). **8** $24.2\ \text{mL}$ (1 dp, limited by the $47.6\ \text{mL}$ reading). **9** (i) $3.5\ \mu\text{g L}^{-1}$; (ii) $3.5 \times 10^{3}\ \text{ng L}^{-1}$ ($= 3500\ \text{ng L}^{-1}$). **10** headings lack units (should be "Temperature / °C" and "Volume of gas / mL"); the entries have inconsistent decimal places (12, 25.0, 48); "Temp" is an abbreviation — use the full quantity name. See solution. **11** (i) evidence; (ii) inference; (iii) opinion. **12** evidence-based rewrite — see solution. **13** The conclusion does not address the aim and is not drawn from the data; correct conclusion: the ethanoic acid concentration was $0.82 \pm 0.03\ \text{mol L}^{-1}$. **14** a No — the counted length is $940 - 180 = 760$ words, above the 600-word maximum. b Incorrect — posters are not limited to a particular paper size or number of panels. c Methodology and methods. d The Conclusion may introduce no new information; the comparison belongs in the Discussion. **15** a one sentence summarising the major finding, at the centre of the poster, occupying 20–25% of the poster space; b the key finding, as a direct answer to the investigation question; c see solution. **16** both must be cited in-text and listed in References; the reproduced diagram must be attributed to its source — see solution. **17** deleting data to fit an expectation is falsification; record all data and exclude only justified outliers transparently — see solution. **18** faults: (1) rates quoted without units; (2) "proves / always / all reactions" over-generalises; (3) the claim ignores uncertainty and is not bounded to the conditions tested. Corrected conclusion — see solution.

Fully-worked solutions

Question 1 a. $3.450\ \text{g}$ has **4** significant figures (the trailing zero after the decimal point counts); $12.0\ \text{mL}$ has **3**; $25.5\ °\text{C}$ has **3**. b. Density $= \dfrac{3.450}{12.0} = 0.2875\ \text{g mL}^{-1}$. The least-precise datum used ($12.0\ \text{mL}$) has 3 significant figures, so the density is quoted as $0.288\ \text{g mL}^{-1}$. c. The **volume**, $12.0\ \text{mL}$, limits the precision because it has the fewest significant figures (3); in division the answer can carry no more significant figures than the least-precise datum. (The thermometer reading is not used in this calculation.)

Question 2 a. $\text{milli} = 10^{-3}$, so $0.0056\ \text{mol L}^{-1} = 0.0056 \times 10^{3}\ \text{mmol L}^{-1} = 5.6\ \text{mmol L}^{-1}$. b. $2500\ \text{mg} = \dfrac{2500}{10^{3}}\ \text{g} = 2.5\ \text{g}$; and $2.5\ \text{g} = \dfrac{2.5}{10^{3}}\ \text{kg} = 2.5 \times 10^{-3}\ \text{kg}$. c. $750\ \text{mL} = \dfrac{750}{10^{3}}\ \text{L} = 0.750\ \text{L}$; and $0.000\,000\,540\ \text{m} = 5.40 \times 10^{-7}\ \text{m} = 540 \times 10^{-9}\ \text{m} = 540\ \text{nm}$ (since $\text{nano} = 10^{-9}$).

Question 3 a. **Evidence** — the statement reports measured data (volumes of gas collected in a fixed time). b. **Inference** — it interprets the evidence, explaining the trend in terms of surface area and reaction sites; it is a reasoned interpretation, not a measurement. c. **Opinion** — it is an unsupported over-generalisation. The data concern only this decomposition; nothing in them shows powders are "best" for "any" experiment, so the claim is a value judgement rather than evidence or a justified inference.

Question 4 a. It is **not** bounded by the data. The investigation measured citric acid content, not "health"; describing Brand C as "healthiest" introduces a claim the evidence does not support. b. It only partly addresses the aim: it correctly identifies Brand C as lowest in acid, but the "healthiest" judgement goes beyond the aim, which concerned acid content. c. A suitable conclusion: *"The three brands differ in citric acid content. Brand A contained the most (*$0.82\ \text{g}$ *per* $100\ \text{mL}$*) and Brand C the least (*$0.55\ \text{g}$ *per* $100\ \text{mL}$*). Brands A and B differ by only* $0.03\ \text{g}$*, which is smaller than the* $\pm 0.04\ \text{g}$ *combined uncertainty of a difference between two measurements each known to* $\pm 0.02\ \text{g}$ *— their ranges,* 0.80–$0.84\ \text{g}$ *and* 0.77–$0.81\ \text{g}$*, overlap — so A and B may not differ at all; Brand C (*0.53–$0.57\ \text{g}$*) is clearly lower than both."* This uses only the data, relates to the aim, and acknowledges the uncertainty.

Question 5 a. **Introduction** — the research question, the aim and the hypothesis set up the investigation. b. **Results** — processed data, including tables of raw titres and labelled graphs such as a calibration curve, are presented here. c. **Discussion** — evaluation of random and systematic errors, and of their effect on the results, is part of interpreting the data.

Question 6 a. The source must be cited **in-text** at the point where the literature value is used, and listed in full in the **References** (author or organisation, year, title, source, and — for an online source — the date accessed), using a standard referencing style. b. The technician's help is recognised in the **Acknowledgements** section of the poster. c. The anomalous titre must be **recorded in the logbook** with all the other raw data. It may be excluded from the mean **only** if there is a justified reason to treat it as an outlier (for example, an identified procedural error, or an objective outlier test), and that exclusion must be **stated and justified** on the poster. The raw value is never deleted or altered, so the data remain authentic.

Question 7 $V_m = \frac{V}{n} = \frac{0.250}{0.0104} = 24.038\ldots \text{ L mol}^{-1}$. Both data have 3 significant figures, so the answer is 24.0 L mol^{-1} (3 sf).

Question 8 Titre $= 47.6 - 23.40 = 24.2 \text{ mL}$. This is a **subtraction**, so the result is limited by the reading with the **fewest decimal places** — the final reading, 47.6 mL, has one decimal place. The titre is therefore quoted as 24.2 mL (1 decimal place), not 24.20 mL; writing an extra decimal place would claim precision the final reading does not have.

Question 9 $3.5 \times 10^{-6} \text{ g L}^{-1}$: (i) $\text{micro} = 10^{-6}$, so $3.5 \times 10^{-6} \text{ g L}^{-1} = 3.5\,\mu\text{g L}^{-1}$. (ii) $\text{nano} = 10^{-9}$, so $3.5 \times 10^{-6} \text{ g L}^{-1} = 3.5 \times 10^{3} \text{ ng L}^{-1} = 3500 \text{ ng L}^{-1}$.

Question 10 Three faults and their corrections: 1. The column headings carry **no units**. They should read "Temperature / °C" and "Volume of gas / mL", so that the entries are pure numbers. 2. The entries have **inconsistent decimal places / significant figures** (12, 25.0, 48). They should be quoted to a consistent resolution matching the measuring instrument, e.g. 12.0, 25.0, 48.0. 3. "Temp" is an informal **abbreviation**; the full quantity name (Temperature) should be used in the heading. (Corrected table: "Temperature / °C" against "Volume of gas / mL", with entries to one decimal place.)

Question 11 (i) **Evidence** — it reports a measured observation and value (the mean titre and the colour change). (ii) **Inference** — it interprets the colour change as the end point and concludes neutralisation occurred; this is reasoning from the evidence, not a raw measurement. (iii) **Opinion** — "easiest to perform" is a personal value judgement unsupported by data.

Question 12 A suitable evidence-based rewrite: *"The results support the hypothesis: adding the catalyst reduced the mean reaction time from* 85 s *to* 32 s*, showing that the reaction rate increased. This conclusion applies only to the reaction and conditions investigated and cannot be generalised to other reactions."* It is quantitative (cites the times), tied to the evidence, uses "support" rather than "prove", and is bounded to the conditions tested.

Question 13 The conclusion does **not** address the aim: the aim asked for a *concentration*, but the statement merely restates prior knowledge (that vinegar is acidic) and is not drawn from the titration data, so it is not evidence-based. A conclusion that does: *"The concentration of ethanoic acid in the vinegar was determined to be* $0.82 \pm 0.03 \text{ mol L}^{-1}$*."* This answers the question directly, uses the data, quotes the uncertainty, and stays within what was measured.

Question 14 a. **No.** The 600-word limit applies to the poster text only: supporting text in tables, graphs, image captions, references and acknowledgements is excluded from the count. The counted length is therefore $940 - 180 = 760 \text{ words}$, which exceeds the maximum, so at least 160 words must be cut. b. The claim is **incorrect**. VCAA states that posters are *not* limited to a particular paper size or number of panels; the requirements are the format, the word limit, and conciseness, clarity and legibility. A single A4 sheet is permitted but not required, and it may well be too small to remain legible. c. **Methodology and methods** — this section summarises the methodology chosen and the data-generation and data-analysis methods, which is where the variables and the procedure are described. d. The Conclusion must respond to the investigation question and state the extent to which the analysis has answered it, **with no new information introduced**; a literature value is new information. Comparing the result with published values is **cross-referencing to chemical concepts**, which belongs in the **Discussion** (with the source cited in References and acknowledgements).

Question 15 a. The **communication statement** is a **single sentence** that summarises the **major finding** of the investigation and answers the investigation question. VCAA places it at the **centre** of the poster, where it occupies **20 –25%** of the poster space. b. A reader who reads nothing else should learn **what the investigation found** — the key

result, expressed as a direct answer to the question. It is *not* an abstract: the aim, method, evaluation and limitations each have their own section and are not repeated in it. c. A suitable communication statement:

"Cooking reduced the vitamin C content of broccoli from 62 mg to 38 mg per 100 g, supporting the hypothesis that cooking lowers vitamin C content."

The essential content is the finding itself ("cooking reduced the vitamin C content of broccoli from 62 mg to 38 mg per 100 g"); the closing clause is acceptable because it keeps the statement to one sentence and ties the finding to the hypothesis.

Question 16 Both borrowed items must be **acknowledged**. For the literature density value: cite the source **in-text** where the value is quoted, and give a full entry in the **References** (author or organisation, year, title, source, date accessed for a website). For the reproduced apparatus diagram: it must be **attributed** to its origin (a caption crediting the source, noting "adapted from…" if changed) and listed in the References; presenting it as the student's own work is **plagiarism**. Following a consistent, standard referencing style throughout is also required.

Question 17 Deleting results **solely because they do not fit the expected trend** is **data falsification** — it misrepresents the evidence to reach a predetermined conclusion, and is scientifically dishonest. The correct procedure is: record **all** raw data in the **logbook** (which authenticates the work) and report them. A result may be set aside only if there is a **justified** reason to treat it as an **outlier** — identified by an objective criterion, not by whether it improves the graph — and that exclusion must be **stated and justified** in the discussion, with the raw value retained in the logbook. All retained data are then used to calculate the mean.

Question 18 Three faults and their corrections: 1. **No units on the rate values** (0.0021, 0.0043) — every quantity must carry its unit, e.g. $\text{mol L}^{-1}\text{s}^{-1}$. 2. **Over-generalisation** — "proves", "always" and "true for all reactions" claim far more than the evidence: a conclusion cannot *prove* a hypothesis, and two data points for one reaction cannot be extended to all reactions or all concentrations. 3. **Ignores uncertainty and is not bounded to the conditions tested** — the data show the rate approximately doubled ($0.0043 / 0.0021 \approx 2.0$) when the concentration doubled, over the narrow range studied, not that it "always" doubles. A corrected conclusion: *"Over the range tested, doubling the concentration from* 0.10 *to* $0.20\,\text{mol L}^{-1}$ *approximately doubled the reaction rate (from* 0.0021 *to* $0.0043\,\text{mol L}^{-1}\text{s}^{-1}$*), which is consistent with the hypothesis that increasing concentration increases the rate for this reaction. Further concentrations would be needed to establish the relationship more fully."*